This volume represents the final output of the interdisciplinary research project CyberPro, funded by the San Raffaele University of Rome. The work was realized through a strategic partnership between the Observatory on Product Cybersecurity based at Università San Raffaele and the Research Centre in European Private Law (ReCEPL) at Università Suor Orsola Benincasa. It brings together revised contributions originally presented during the project's final international conference, which took place on 29 November 2024 at the Italian Supreme Court. The collection of academic essays explores the complex link between cybersecurity, technological innovation, and legal liability within interconnected societies. The contributions analyse the impact of crucial European regulations, starting from the AI Act, the NIS 2 Directive, and the new discipline on liability for defective products, applying them to critical sectors such as digital healthcare and autonomous mobility. Particular attention is paid to the protection of vulnerable subjects and the management of risks arising from algorithmic systems, proposing a shift from a reactive to a preventive model of liability. The text furthermore examines advanced technological tools, including blockchain and ISO certifications, as means to ensure transparency and data security. In this light the work delineates the new frontiers of European private and public law in an attempt to harmonize digital progress with the safeguarding of fundamental rights and human dignity.

Giovanna Capilli is Full Professor of Private Law and Director of the Cybersecurity Observatory of Products at San Raffaele University Rome.

ISBN 979-12-5511-031-6

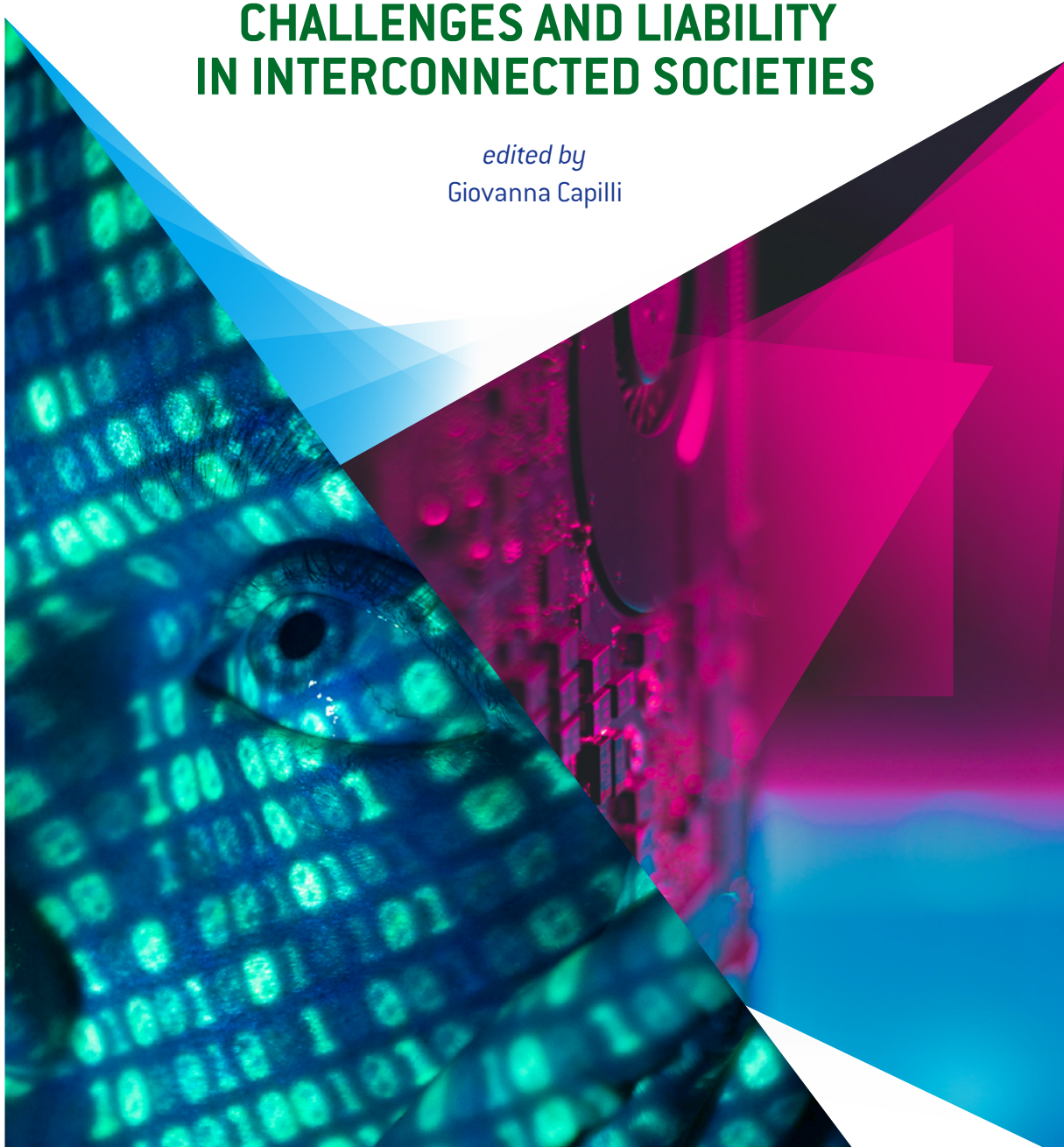
ReCEPL SERIES 12 CYBERSECURITY AND PRODUCTS: CHALLENGES AND LIABILITY IN INTERCONNECTED SOCIETIES

ReCEPL SERIES
12.

SUOR ORSOLA
BENINCASA
UNIVERSITÀ EDITRICE

CYBERSECURITY AND PRODUCTS: CHALLENGES AND LIABILITY IN INTERCONNECTED SOCIETIES

edited by
Giovanna Capilli




UNIDIR
ReCEPL
4STAI
Jean Monnet
Centre of Excellence



ReCEPL SERIES

1. **Gatt L., Montanari R., Caggiano I.A.** (*eds.*)
Privacy and Consent. A Legal and UX&HMI Approach University Suor Orsola Press, 2021,
ISBN 978-88-96055-878. https://doi.org/10.57230/RS1

2. **Gatt L.** (*ed.*)
Social Networks and Multimedia Habitats. Jean Monnet Chair PROTECH.
"European Protection Law of Individuals in Relation to New Technologies".
1st International Workshop A.Y. 2019-2020 University Suor Orsola Press, 2020,
ISBN 979-12-80426-00-0. https://doi.org/10.57230/RS2

3. **Grimaldi P.**
Le fonti di energie rinnovabili nella circular economy University Suor Orsola Press, 2021,
ISBN 979-12-80426-05-5. https://doi.org/10.57230/RS3

4. **Gatt L.** (*ed.*)
Legal design for transparency in consumer contracts operating in online and offline environments.
A new chance for the interaction between academics and stakeholders.
2nd International Workshop A.Y. 2020-2021 University Suor Orsola Press, 2022,
ISBN 979-12-80426-10-9. https://doi.org/10.57230/RS4

5. **Gatt L., Gaeta M.C., Aulino L.** (*eds.*)
Dall'essere umano al cyborg. I casi nazionali sugli embrioni umani e la Corte di giustizia dell'UE: dalla procreazione artificiale al potenziamento umano nell'era del transumanesimo. Jean Monnet Chair PROTECH. "Il diritto europeo a tutela della persona in relazione alle nuove tecnologie",
3° Workshop Internazionale AA 2021-2022, Suor Orsola Università Editrice, 2023,
ISBN 979-12-5511-003-3. https://doi.org/10.57230/RS5

6. **Caggiano I.A.** (*ed.*)
Design of rights for eco-food and eco-fertility: markets and causal relations. Jean Monnet Chair EUGREENEXT "European Green Rights: reshaping fundamental rights for next generations"
1st International Workshop A.Y. 2021-2022, Suor Orsola Università Editrice, 2023,
ISBN 979-12-5511-007-1. https://doi.org/10.57230/RS6

7. **Grimaldi P.**
End of waste nell'attività di impresa tra sostenibilità, artificial intelligence e digital innovation,
Suor Orsola Benincasa Università editrice, 2024, ISBN 979-12-5511-013-2. https://doi.org/10.57230/RS7

8. Caggiano I.A., Mollo A.A. (*eds*.)
EU Green Rights: re-shaping fundamental rights for future generations.
La sostenibilità nella prospettiva europea dei diritti, Suor Orsola Università Editrice, 2025,
ISBN 979-12-5511-016-3. https://doi.org/10.57230/RS8

9. Aulino L.
Informazione giuridica, volontà dell'utente e legal design nei contesti di automazione. Il caso del privacy disclaimer nel veicolo a guida autonoma, Suor Orsola Università Editrice, 2025,
ISBN 979-12-5511-022-4. https://doi.org/10.57230/RS9

10. Mollo A.A.
Neurotecnologie e soggetti vulnerabili, Suor Orsola Università Editrice, 2025, ISBN 979-12-5511-023-1.
https://doi.org/10.57230/RS10

11. A cura di **A.A. Carrabba, L. Gatt, G. Iaccarino, con la collaborazione di Irene Coppola, Luigi Izzo**
Questioni di diritto successorio in ricordo di Giancarlo Laurini, Suor Orsola Università Editrice,
2026, ISBN 979-12-5511-030-9. https://doi.org/10.57230/RS11

12. G. Capilli *(ed)*
Cybersecurity and Products: Challenges and Liability in Interconnected Societies,
Suor Orsola Università Editrice, 2026, ISBN 979-12-5511-031-6. https://doi.org/10.57230/RS12

CYBERSECURITY AND PRODUCTS: CHALLENGES AND LIABILITY IN INTERCONNECTED SOCIETIES

edited by
Giovanna Capilli

Università San Raffaele Roma

"La pubblicazione è stata finanziata dal progetto CyberPro dell'Università Telematica San Raffaele Roma con i fondi del finanziato competitivo FIN/RIC 2023"

SUOR ORSOLA
BENINCASA
UNIVERSITÀ EDITRICE

La ReCEPL4STAI Series rappresenta una implentazione dell'omonimo Centro di eccellenza Jean Monnet.
La serie ReCEPL4STAI è diretta da Lucilla Gatt.
I membri del comitato scientifico sono disponibili qui:
https://www.unisob.na.it/ateneo/c008_u.htm?vr=1

Questa ricerca è stata sottoposta a revisione anonima.

Editing
Luciana Trama

Copertina
Flavia Soprani

Progetto grafico e impaginazione
Anna Orefici RE:LAB

Stampa
Carmine Marra

© Copyright 2026 dell'Università Suor Orsola Benincasa. Tutti i diritti sono riservati.
I singoli saggi rimangono di proprietà intellettuale degli autori.

ISBN 979-12-5511-031-6

Il libro è rilasciato sotto licenza Creative Common Attribution 4.0 International Licence CC-BY-NC-ND.
Tutti i dettagli sulla licenza sono disponibili all'indirizzo: creativecommons.org/licenses/by- nc-nd/4.0.
Accesso gratuito al libro su: *www.unisob.na.it/ateneo/c008_u.htm?vr=1*.
La proprietà intellettuale dei singoli contributi rimane degli autori.

SOMMARIO

Lucilla Gatt
Project Presentation **15**

Giovanna Capilli
Introduction **17**

CAPITOLO 1
CYBERSECURITY OBLIGATIONS AND ISO CERTIFICATIONS TOWARDS A NEW MODEL OF LIABILITY FOCUSED ON DAMAGE PREVENTION **21**
Lucilla Gatt
1. Digital evolution and the new European regulatory horizon 21
2. Sanctions regime and interaction with civil liability 22
3. Manufacturer liability and IT defects 23
4. Harmonisation between positive law and international technical standards 25
5. Towards a model of mandatory pre-market ISO certification for damage prevention 27

CAPITOLO 2
AI ACT AND AUTOMATED DRIVING BETWEEN CURRENT DEVELOPMENTS AND EXPERIMENTATION **29**
Enrico Al Mureden
1. Premises 30
2. The five levels of driving automation between technical feasibility and regulatory transformation 31
3. The "Vision Zero" objective and the containment of human error between Advanced Driver Assistance Systems (ADAS) and driverless cars 33

4. Autonomous driving between regulatory sandboxes and real-world testing 35
5. Automated driving and new frontiers of civil liability 38

CAPITOLO 3
CYBER SECURITY AND HEALTHCARE FACILITIES **45**
Antonina Astone
1. Introduction 45
2. The social function performed by data and the measures for their safe treatment 47
3. The liability of administrators foreseen by the NIS 2 Directive 49
4. The possible alternative offered by blockchain 50

CAPITOLO 4
CYBERSECURITY AND THE PROTECTION OF VULNERABLE GROUPS IN THE DIGITAL AGE: CHALLENGES AND PREVENTION STRATEGIES **55**
Irene Coppola, Cristina Mingo
1. Introduction 56
2. Cybersecurity challenges in organisations 56
3. The digital jungle: online risks 57
4. Regulatory developments in online protection 58
5. The role of Digital Services Act (DSA) 59
6. Artificial Intelligence and the protection of children online 59
7. The importance of Cybersecurity education: the case of Iris Stein 60
8. Conclusions and future prospects: legal design for the teaching of private law 60

CAPITOLO 5
CYBERSECURITY AND MEDICAL DEVICES: A PROGRESSIVE APPROACH IN DIGITAL BIOETHICS **63**
Giuseppe Corasaniti
1. Introduction and Foundations of Digital Bioethics 64
2. EU Regulatory Framework: MDR, GDPR, NIS2, Cybersecurity Act 66
3. CJEU Jurisprudence: Proportionality, Liability, and AI Oversight 67
4. ECHR Article 2: Right to Life and State Obligations for Medical Technologies 71
5. Digital Bioethics Applied: Codes of Conduct and Ethical Governance 77

CAPITOLO 6
TECHNOLOGICAL INNOVATION, VULNERABILITY AND AGEING.
THE PROTECTION OF THE INDIVIDUAL **81**
Stefano Corso
1. Law and technological innovation. Vulnerability of technology and vulnerability of the subject 82
2. The elderly and the digital 85
3. Safety, security and risk. A specific problem: algorithmic discrimination 88
4. Product Cybersecurity and protection of the individual 92

CAPITOLO 7
THE AGENTIC PARADOX: AN ACADEMIC ANALYSIS OF EMERGENT CYBERSECURITY THREATS IN "STATE-OF-THE-ART" AI SYSTEMS **99**
Gerardo Costabile
1. Introduction: the dawn of the "agentic era" 100
2. A paradigm shift in web architecture and security 101
   2.1 The erosion of traditional security boundaries 101
   2.2 The Model Context Protocol (MCP) and the Rise of Dynamic Integration 102
3. A new taxonomy of threats in agentic ecosystems 102
   3.1 Data-driven attacks: content injection and exfiltration 103
   3.2 Supply chain attacks in a decentralized AI ecosystem 103
   3.3 The Agent as an inadvertent adversary 104
4. Industry response and current mitigation postures 104
   4.1 The precautionary principle: Gartner's recommendation to block AI browsers 105
   4.2 Technical implementation of blocking policies 105
   4.3 The inadequacy of legacy data loss prevention (DLP) 106
5. Towards a resilient security framework for Agentic AI 106
   5.1 A defense-in-depth control framework 106
   5.2 The MCP gateway architecture: a centralized enforcement point 107
   5.3 Aligning with governance frameworks and defining future research 108
6. Conclusion 108

CAPITOLO 8
CYBERSECURITY AND SOME CHALLENGES OF ENCRYPTION AND PRIVACY **111**
Isadora de Morais Maria, Rafael de Brito Dias,
Alcides Eduardo dos Reis Peron
1. Technovigilance and AI in the contemporary age 112
2. Cybersecurity and some challenges of encryption and privacy 114
3. A dialogue between cryptography and Cybersecurity 116
4. A dialogue between privacy and Cybersecurity 117
5. Cybersecurity and its challenges: the case of China 118
6. Cybersecurity and its challenges: the case of Brazil 120
7. Conclusion 122

CAPITOLO 9
AI AND TORT **127**
Luca Di Donna
1. Defining a System of Civil Liability Applicable to Artificial Intelligence 128
2. Prospects for Reform in Balancing Technological Development and the Protection of Fundamental Rights 136

CAPITOLO 10
ARTIFICIAL INTELLIGENCE AND CYBERSECURITY: EUROPEAN LAW AND REGULATORY TOOLS **143**
Fernanda Faini
1. Legal rules in the digital society 144
2. The new face of law: changes in the form of regulation 148
3. The innovative content of the law: paradigms, instruments and solutions 152
   3.1 The evolution of the relationship between law and technology: technological humanism, incorporation of law into technology, and the risk-based approach 152
   3.2 Human in the loop: human supervision and surveillance 155
   3.3 Algorithmic transparency: explainability of Artificial Intelligence and contestability of technology 156
4. Conclusions: philosophical and legal perspectives and scenarios 158

CAPITOLO 11
RISK ASSESSMENT AND MANAGEMENT OF ARTIFICIAL INTELLIGENCE SYSTEMS: THE CASE STUDY OF BRAIN COMPUTER INTERFACE DEVICES **165**
Maria Cristina Gaeta
1. The concept of 'risk' in the regulation of Artificial Intelligence 166

2. Brain-computer interface devices as an example of high-risk AI 167
3. Institutional interventions to guarantee safe and reliable neurotechnologies 169
4. Risk assessment and management of BCI *technologies*: the role of standardisation 172

CAPITOLO 12
CYBERSECURITY AND CRIMINAL LAW: ASPECTS OF CORPORATE **183**
Andrea Gentile
1. Introduction 184
2. Cybersecurity and corporate criminal liability 186
3. D.L. Cybersecurity and the impact on corporate crime liability 187
4. Critical considerations and future perspectives 189

CAPITOLO 13
RELATIONSHIPS AND INTERCONNECTIONS BETWEEN CYBERSECURITY AND CORPORATE LIABILITY: THE "DERIVED" OBLIGATION OF CRIME RISK PREVENTION **191**
Patrizia Giusti
1. Introduction 192
2. New Cybercrime law (DDL) 192
3. The new features of the law 193
4. The new features of the law 194
5. Extortion through computer crime 196
6. The Method 197
7. Conclusions 198

CAPITOLO 14
THE ALGOMORPHIC SOCIETY: TOWARDS A THEORY OF ALGORITHMICALLY MEDIATED IDENTITIES DEVICES **201**
Edmondo Grassi
1. From the Digital Self to the Algomorphic Self 202
2. An Identity Morphogenesis 203
3. Reframing the Social Beyond Anthropocentrism 204

CAPITOLO 15
CHILD SAFETY ONLINE: HOW AI CAN BE USEFUL AND WHAT ARE THE CRITICAL ISSUES? **207**
Michele Iaselli
1. AI's Contribution 208
2. Critical Issues Related to Ai Use 210

3. The AI Act Contribution 213
4. Responsabilities Arising from AI Use 215

CAPITOLO 16
BLOCKCHAIN AT THE SERVICE OF CYBERSECURITY IN THE ITALIAN FOOD AND AGRI-FOOD SUPPLY CHAIN **219**
Enrichetta Giulia Jannotti Pecci
1. Introduction 220
2. Why blockchain? 223
3. The use of blockchain to guarantee traceability and safety of the food and agri-food supply chain 227
4. Protecting the agricultural and agri-food supply chain and intellectual property from the risks of data breaches 229
5. Conclusions 231

CAPITOLO 17
CYBERSECURITY IN LAW FIRMS BETWEEN DATA PROTECTION AND CUSTOMISATION REQUIREMENTS **235**
Valeria Manzo
1. The use of technology in the practice of law 236
2. The principles of the GDPR 239
3. The IBM 2023 Security report 241
4. So-called 'common' attacks in the legal sector and the importance of Cybersecurity training 243
5. Final considerations 249

CAPITOLO 18
THE RELATIONSHIP BETWEEN GDPR AND BLOCKCHAIN: BETWEEN DATA PRIVACY AND TECHNOLOGICAL SECURITY **253**
Federica Marani
1. Introduction. The processing of personal data 254
2. The "new technologies": smart contracts and blockchain 255
   2.1 Blockchain permissionless and permissioned, public and private: differences and implications **258**
3. Is it possible to find a connection between GDPR and blockchain technology? 260
4. Liability and Compensation profiles 262

CAPITOLO 19
ARTIFICIAL INTELLIGENCE: BETWEEN MARKET, REGULATION, SUBJECTIVITY AND RISK ALLOCATION OF RESPONSIBILITY **267**
Giuseppe Maria Marsico
1. Introductory remarks 268
2. Allocation of liability risk and the market: do we really need to rethink dogmatic categories of subjectivity and liability? 271
3. Brief concluding remarks: between risk-based approach of the AI and balancing of interests 275

CAPITOLO 20
JURISPRUDENCE AND ARTIFICIAL INTELLIGENCE THE ROMEO AND JULIET OF LAW **283**
Rachele Ranieri
1. The die is cast: a first approach to predictive justice 284
2. What's good 285
3. The oder side of the coin 287
4. Beyond good and evil 294

CAPITOLO 21
THE ACCOUNTABILITY AND LIABILITY PARADIGMS APPLIED TO AI: THE EUROPEAN FRAMEWORK **299**
Simone Saccomani
1. Introduction 299
2. The principles of accountability and liability applied to AI 301
3. AI civil liability as an autonomous compensation case 304
4. The European legal framework: the AI Liability Directive and the new Product Liability Directive 307
5. Conclusions 311

CAPITOLO 22
MINORS AND AGE VERIFICATION SYSTEMS: THE PROTECTION OF THE «MOBILE BORN» IN THE INFORMATION SOCIETY **317**
Edoardo Salvatore
1. The child as a vulnerable subject of the network 318
2. The protection of minors in the European landscape: reflections in the light of the GDPR 321
3. Age verification systems: a critical analysis of evolving concepts 324

CAPITOLO 23
NIS 2: THE EUROPEAN STRATEGY ON CYBERSECURITY
BETWEEN NOTIFICATION OBLIGATIONS AND IMPACT ON STAKEHOLDERS **331**
Viviana Visaggio
1. The European Cybersecurity strategy 333
2. The NIS 2 Directive 336
3. Entities subject to NIS 2 obligations: sectors and parameters 340
4. The supply chain and the importance of Cybersecurity: obligations and responsabilities for system resilience 342
5. Notification obligations in the event of cyber incidents 344
6. Penalities and publication obligations 345
7. A reflection on the implications for those involved 347

CAPITOLO 24
THE "DIGITAL MINORS" BETWEEN VULNERABILITIES AND OPPORTUNITIES
FOR RISK CONTAINMENT. **351**
Lavinia Vizzoni
1. "Digital Minors" Between Risks and Opportunities 352
2. Minors and Artificial Intelligence Act 355
3. An understimated risk: the use of video games by minors 357
4. A first conclusion: the role of certifications and Cybersecurity 361

LUCILLA GATT

Project Presentation

This volume forms part of the research project CyberPro (FIN/RIC) 2023, funded by the San Raffaele University of Rome. The project is interdisciplinary in nature and is directed by Professor Giovanna Capilli ReCEPL4STAI key member, is realaised with the support of the Jean Monnet Centre of Excellence ReCEPL4STAI based at Suor Orsola University. The research team includes Professors Marina D'Amato, Edmondo Grassi, Andrea Gentile, Antonio Pallotti, and Noemi Scarpato. The project has been developed through the complementary and integrated contributions of private law, computer science, criminal law, and sociology.

The investigation is explicitly situated within a de iure condendo perspective and seeks to analyse the impact that new regulatory provisions and their related implementation measures may have on the existing legal framework, understood in its dynamic dimension and in its constant tension between stability and innovation. Within this framework, particular attention is devoted, on the one hand, to the identification of cyber and Cybersecurity risks affecting specific categories of products and, on the other, to issues concerning the allocation of liability in the event of damage, as well as to the legal treatment of products incorporating Artificial Intelligence systems.

The volume brings together revised and updated contributions presented at the conference held on 29 November 2024, which represented a pivotal moment from Italian and foreign researchers, including those affiliated with ReCEPL4STAI, of collective discussion and scholarly exchange during the research project. That debate constituted a privileged forum for dialogue among different disciplinary and methodological approaches, allowing — through a critical and non-assertive lens — for the identification of possible trajectories of regulatory evolution and for an assessment of their systemic implications, in line with the de iure condendo approach that permeates the entire work.

In accordance with the objectives of the University's Third Mission, the

research activities culminated, in January 2025, in the establishment of an Observatory on Product Cybersecurity, under the direction of Professor Giovanna Capilli. The Observatory is conceived as a hub for the convergence of heterogeneous professional expertise, for continuous dialogue with relevant stakeholders, and for contributing to risk analysis and to the development of analytical tools and operational protocols useful to economic operators and national authorities throughout the phases of product design, assessment, and market placement, also with a view to the protection of consumers and users. Among the Observatory's main tasks are: monitoring legislative and judicial developments; creating databases relating to initiatives addressing Cybersecurity challenges; identifying best practices, control systems, and certification models; conducting research activities and issuing opinions for enterprises and institutions; as well as overseeing compliance with obligations concerning transparency, updating, and vulnerability management, which today constitute essential prerequisites for a form of legal liability adequate to the transformations of markets and technology.

GIOVANNA CAPILLI*

Introduction§

Legal reflection on Cybersecurity and digital products today lies at the core of a broader process of redefinition of the fundamental categories of contemporary law. The increasing pervasiveness of information technologies, the systematic use of decision-making algorithms, and the integration of Artificial Intelligence into essential services affect not only economic and social dynamics, but also the very structure of the legal safeguards designed to protect individuals.

In this context, Cybersecurity cannot be regarded as a merely technical or instrumental dimension. Rather, it emerges as a necessary precondition for the effective enjoyment of fundamental rights and as a privileged domain in which new forms of responsibility borne by manufacturers, digital service providers, and, more broadly, actors operating within the technological ecosystem come to the fore. Law is thus called upon to confront unprecedented forms of risk, which no longer arise solely from human conduct, but also from the — often opaque — functioning of automated systems and complex digital products.

* Full Professor of Private Law at the San Raffaele University of Rome.

§ This contribution develops the presentation at the conference "Cybersecurity e prodotti: sfide e responsabilità nelle società interconnesse", 29 November 2024, Italian Supreme Court.

The Editor would like to express her sincere gratitude to Maria Cristina Gaeta, Paola Giacalone, Sergio Perrotta and Ludovica Maria Giordano for their valuable support in the preparation of this volume.

This volume is grounded in this awareness and seeks to offer a systematic contribution to the understanding of the evolving contours of legal regulation in the field of Cybersecurity, with particular reference to digital products and the repercussions their use may have on individuals. The essays collected herein follow a shared trajectory: an analysis of how law — especially at the European level — is progressively developing conceptual and normative tools capable of governing technological innovation without relinquishing its protective function.

A first element that clearly emerges is the ambivalent nature of technology. Technology represents, at once, a driver of development and a potential multiplier of inequalities and risks. Digital products can enhance individual autonomy, improve access to services, and raise quality of life; however, if inadequately designed or regulated, they may also expose individuals to new forms of discrimination, exclusion, and infringement of rights. Cybersecurity is situated precisely within this tension, functioning as a balancing criterion between innovation and protection and constituting an essential prerequisite throughout the entire life cycle of Artificial Intelligence systems and models.

Within this framework, the concept of vulnerability assumes a central role and has been progressively articulated in an increasingly sophisticated manner within European Union law. Alongside technological vulnerability — understood as an intrinsic weakness of a product or IT system susceptible to exploitation by external threats — the notion of human vulnerability has emerged, referring to personal conditions that render certain individuals more exposed to digital risks. While this distinction is necessary, it does not exhaust the complexity of the phenomenon, since the interaction between individuals and technology generates forms of vulnerability that do not coincide either with technical defects or with individual fragility, but rather arise from the relationship itself with increasingly complex and automated systems.

The contributions collected in this volume devote particular attention to these dynamics with reference to older persons and minors. Age is not treated as a rigid legal category, nor as an automatic presumption of weakness, but as a possible risk factor that law must address in a differentiated manner. In areas such as digital healthcare, personal data protection, and access to online services, the use of insecure or non-inclusive technological products may produce significant detrimental effects, highlighting the close connection between Cybersecurity, product design, and the protection of the individual.

From this perspective, the volume underscores the central role played by the European legislator in constructing a regulatory model based on a risk-based and person-oriented approach. The GDPR (EU) 2016/679, the Artificial Intelligence

Act (EU) 2024/1689, the NIS 2 Directive, the Cyber Resilience Act (EU) 2024/2847, the Digital Services Act (EU) 2022/2065, the Data Act (EU) 2023/2854, and Directive (EU) 2024/2853 on liability for defective products outline a normative framework that does not merely intervene ex post, but rather permeates the entire life cycle of digital products, imposing security obligations, ex ante impact assessments, and accountability mechanisms on economic and organisational operators. Cybersecurity thus becomes a structural requirement of legal compliance, embedded within decision-making and production processes.

Finally, the proposed "Digital Omnibus" Regulation (COM(2025) 837), presented by the European Commission on 19 November 2025, constitutes a structural intervention in the Union's digital regulatory architecture as a whole. Its aim is to consolidate into a single normative corpus provisions currently dispersed across multiple legislative acts, pursuing the dual objective of ensuring systemic coherence within the European data acquis and reducing administrative burdens for businesses and public administrations, while at the same time maintaining a high level of protection in the areas of data, privacy, and Cybersecurity.

Alongside the European dimension, the volume highlights the contribution of national legal systems, and in particular the Italian experience, which develops in constant dialogue with EU law. Initiatives aimed at promoting digital inclusion, digital literacy, and reducing of generational divides demonstrate how legal protection cannot be detached from active policies capable of accompanying technological innovation.

Taken as a whole, the essays collected herein seek to provide a systematic reading of a rapidly evolving field, in which the Cybersecurity of digital products emerges as one of the new frontiers of European private and public law — a frontier against which the capacity of law to preserve its ordering function is tested, ensuring that technological development remains consistent with the fundamental values of the person, dignity, and solidarity that underpin the legal order.

LUCILLA GATT*

Cybersecurity obligations and ISO certifications Towards a new model of liability focused on damage prevention§

1. Digital evolution and the new European regulatory horizon

The progressive integration of Artificial Intelligence and evolutionary software into consumer products has led to the traditional concept of goods as static entities being superseded, ushering in what legal doctrine defines as the 'agentic turn', i.e. a phase characterised by the ability of digital systems to act autonomously on the physical and information environment. In this context, Cybersecurity is no longer a mere technical attribute, but an essential legal requirement for the protection of fundamental rights, consumer safety and the integrity of the internal market.

The European Union has responded to this transformation by building a multi-level and systemic regulatory framework based on directly applicable regulations and enhanced harmonisation directives.

In particular, the AI Act (Reg. EU 2024/1689) introduces a classification of Artificial Intelligence systems based on the risk of impact on fundamental human rights, imposing stringent governance, transparency and security requirements for high-risk applications; the CRA-Cyber Resilience Act (EU Reg. 2024/2847) establishes horizontal Cybersecurity requirements for all products with digital elements, throughout their entire life cycle; the NIS 2 Directive (EU Dir. 2022/2555) strengthens risk management and incident reporting obligations, directly affecting the liability of administrative bodies. However, a decisive step in the evolution of European law towards consumer protection, also from the perspective of product Cybersecurity, is represented by the new

* Full Professor of Private Law and New Technologies Law at Università degli Studi Suor Orsola Benincasa. Director of the Research Centre in European Private Law (ReCEPL).

§ This contribution develops the presentation at the conference 'Cybersecurity e prodotti: sfide e responsabilità nelle società interconnesse', 29 November 2024, Italian Supreme Court.

Directive on liability for defective products (PLD-Directive EU 2024/2853), which definitively overcomes the dichotomy between tangible and intangible products by expressly including software, AI systems, updates and digital services functionally integrated into the product. The notion of defect is thus recalibrated in light of legitimate consumer safety expectations, taking into account, among other things, the Cybersecurity requirements provided for by European law and, in particular, by the aforementioned legislative acts.

The result is a regulatory ecosystem in which Cybersecurity takes on the role of a structural condition for the legitimacy of placing a product on the European market, according to a logic of ex ante compliance aimed at preventing damage.

2. Sanctions regime and interaction with civil liability

Failure to comply with the requirements of the CRA-Cyber Resilience Act entails severe consequences ranging from fines running into millions to physical exclusion from the European market and increased civil liability in the event of damage to consumers.

The CRA's penalty framework is extremely rigorous and provides, first and foremost, for administrative penalties proportionate to the seriousness of the violation and the company's turnover. In the event of a breach of the fundamental obligations of secure design and development ('security by design'), the manufacturer risks penalties of up to £15 million or 2.5% of the previous year's global turnover, whichever is higher. For formal or CE marking-related non-compliance, fines can reach €10 million or 2% of global annual turnover. Similarly, providing incorrect or misleading information to the authorities can result in penalties of up to €5 million or 1% of turnover.

In addition to monetary penalties, supervisory authorities have powerful powers to protect the integrity of the single market. A product that does not meet the requirements can be immediately banned from the EU market, and authorities can order the withdrawal or recall of products already on the market if they present significant vulnerabilities. The CRA also requires actively exploited vulnerabilities and serious incidents to be reported within very tight time frames (24-72 hours). Failure to notify constitutes an independent punishable violation. Finally, manufacturers are legally obliged to provide security updates for at least 5 years (or for the expected lifetime of the product). Early termination of support constitutes a violation of the Regulation.

Italy is implementing these provisions through the European Delegation Law 2025 (Article 16), which tasks the Government with establishing the

CRA and defining a national penalty system that is effective, proportionate and dissuasive, coordinating it with the existing regulations of the National Cyber Security Perimeter and Legislative Decree 138/2024 (NIS 2), and also to identify the National Cybersecurity Agency (ACN) as the market notification and supervisory authority, with the power to impose sanctions and corrective measures.

In summary, in addition to financial penalties, the supervisory authorities have prohibitive and corrective powers, including the prohibition of marketing, withdrawal or recall of non-compliant products.

However, one of the most significant consequences stems from the intersection of the CRA with the new Directive on liability for defective products (EU Directive 2024/2853-PLD). This interaction further amplifies the manufacturer's exposure: failure to comply with Cybersecurity requirements not only legitimises repressive intervention by the administrative authorities, but also determines a presumption of product defectiveness in civil proceedings, with a reversal of the burden of proof and a consequent strengthening of compensation protection (see paragraph 3).

In conclusion, manufacturers who ignore the CRA risk not only economic loss, but also commercial paralysis in the EU and a position of extreme weakness in any civil dispute brought by consumers.

3. Manufacturer liability and IT defects

The regulation of Cybersecurity defects in products placed on the market has undergone, as mentioned, a profound evolution with the introduction of Directive (EU) 2024/2853 (PLD), which incorporates the provisions of the Cyber Resilience Act (CRA-EU Regulation 2024/2847) and the AI Act (EU Regulation 2024/1689) to ensure consumer protection in the digital age. According to the combined reading of these acts, a product is considered defective if it does not offer the safety that a consumer can legitimately expect, taking into account all circumstances, including Cybersecurity requirements relevant to physical and data security.

More specifically, the Product Liability Directive (PLD) establishes a close link between compliance with Cybersecurity standards and the civil liability of the manufacturer. In particular, Article 10 of the PLD introduces a presumption of product defectiveness if the claimant proves non-compliance with mandatory safety requirements established by Union or national law. This explicit reference to product safety must be understood in a broad sense,

i.e. including the specific form of safety that is Cybersecurity/Cybersecurity as a particular type of safety relating to all products operating in the digital sphere: on this point, PLD, Article 7(2)(f) expressly states that 'in assessing the defectiveness of a product, all circumstances shall be taken into account, including the relevant product safety requirements, including Cybersecurity requirements relevant to safety'. And even earlier, Recital 32 states that 'a product may also be defective due to its vulnerability in terms of Cybersecurity, for example if it does not meet the Cybersecurity requirements relevant to safety'. Similar indications are also clearly written in Recitals 34, 51 and 55 of the same Directive.

This means that (as of the entry into force of the PLD) an unresolved IT vulnerability (in terms of Cybersecurity) or failure to provide updates necessary to maintain the (cyber) security of the product may directly constitute the manufacturer's strict liability. The manufacturer remains liable for defects that occur even after the product has been placed on the market, if it retains control over the software or related services through updates or machine learning algorithms.

As mentioned, Article 10 of the PLD introduces a legal presumption of defect if the injured party demonstrates non-compliance with regulatory (Cyber) security obligations, including those arising from the CRA or the AI Act, significantly reducing the burden of proof on the consumer. Cyber vulnerability, failure to implement security and Cybersecurity measures by design, or failure to provide the security updates necessary to address emerging risks therefore constitute an independent form of product defect.

Particularly relevant is the extension of the manufacturer's liability to the post-market phase: where the manufacturer retains control over the product through software updates, machine learning systems or remote services ( ). In such cases, the manufacturer is also liable for defects that arise after the product has been placed on the market. From this perspective, product liability is a dynamic liability, closely linked to the technological governance of the product's life cycle.

In other words, it can be said that the new PLD Directive has established:

a) an expansion of the concept of 'product' and 'defect', formally equating software with physical products and including operating systems, apps, firmware and Artificial Intelligence systems. A product may be declared defective specifically because of its vulnerability in terms of Cybersecurity (cyber vulnerability), if this weakness exposes the user to risks to health, property or data integrity. The assessment of defectiveness must also consider

the effect that other products (such as smart home devices) may have on the product in question through interconnection;

b) liability for the entire product life cycle (post-market): the manufacturer remains liable for defects that occur after the product has been placed on the market if it retains control over the product through software updates or machine learning algorithms (manufacturer control). Failure to provide security and Cybersecurity updates necessary to remedy emerging vulnerabilities is explicitly defined as a defect that prevents the manufacturer from being exempt from liability (lack of updates). The manufacturer is generally liable for 10 years, but this period is extended to 25 years for personal injury that manifests itself late (term of liability);

c) legal presumptions and a reversal of the burden of proof: the defective nature of a product is legally presumed if the claimant demonstrates non-compliance with mandatory safety requirements established by Union law (such as those provided for in the Cyber Resilience Act or the AI Act). A presumption of product defectiveness is therefore established in certain circumstances. Furthermore, where technical or scientific complexity (typical of AI systems) makes it excessively difficult for the consumer to prove the defect, the courts may presume defectiveness or causality if the damage is probable (technical difficulties-black box). Finally, manufacturers are obliged to disclose relevant evidence in court; in case of refusal, the defect is automatically presumed by the judge (obligation of disclosure of evidence by manufacturers).

4. Harmonisation between positive law and international technical standards

In order for the principles set out by the European legislator to be implemented effectively and efficiently, it is essential that the regulatory framework be harmonised with international technical standards, in particular with the standards of the ISO (International Organisation for Standardisation), IEC (International Electrotechnical Commission) and UNI (Italian National Unification Body). And, in fact, while the European legislator defines 'essential objectives' of security, reliability and resilience, international technical standards provide the operational specifications needed to translate these objectives into verifiable and measurable practices.

The CRA (Cyber Resilience Act) is a regulation that imposes horizontal Cybersecurity requirements, obliging manufacturers to guarantee 'security by design' and to report actively exploited vulnerabilities within 24-72 hours, as

already mentioned. In this context, the implementation of technical standards, such as ISO/IEC 27402:2023-IoT security and privacy and the ISO/IEC 27000 series ("Information Security Management Systems (ISMS) Family of Standards") for information security, constitute the essential technical basis for ensuring the 'security by design' required by European regulations.

However, technical standards also provide the operational specifications necessary to implement, for example, human oversight, robustness and accuracy of AI systems. The ISO/IEC 42001:2023 standard – 'Information Technology – Artificial Intelligence – Management System' – is emerging as the tool of choice for AI governance, anticipating – in chronological order – how many of the legal requirements of the AI Act (approved subsequently) will be implemented in concrete and measurable management practices.

ISO/IEC 42001:2023 has recently been joined (following the adoption of the AI Act) and in a relationship of mutual integration, by ISO/IEC 42005:2025-'Information Technology-Artificial Intelligence (AI) – AI system impact assessment', which provides guidelines for organisations that need to conduct impact assessments of Artificial Intelligence (AI) systems, helping them to identify, analyse and manage the effects (positive and negative) of an AI system on individuals, groups and society throughout its entire life cycle.

ISO/IEC 42005:2025 is a practical standard for risk assessment and the promotion of responsible, ethical and transparent AI, applicable to any organisation that develops, supplies or uses AI systems, regardless of size or sector. It is aimed at all organisations that create, supply or – note – even just use Artificial Intelligence systems, from large technology companies to SMEs, and has a number of key features: a) it covers the entire AI life cycle, from design to withdrawal; b) it is a guide, not a certification standard, but integrates with ISO/IEC 42001: 2023 on AI management systems; c) it suggests how to integrate impact assessment into risk management and existing management systems; d) it defines steps such as documentation, assignment of responsibilities, definition of thresholds (triggers) for assessments and continuous monitoring. The main purpose of ISO/IEC 42005:2025 is to establish a framework for assessing how AI systems and their foreseeable applications may affect people and society, as well as to promote trust and accountability by identifying potential impacts (including unintended ones such as misuse). In summary, ISO 42005 helps to understand 'who could be affected', 'when to assess' and 'what constitutes a significant impact', making AI systems safer and more reliable.

In this context, the time has come to clearly question the role of technical standards (ISO/IEC/UNI) in the reasoning process that leads to the at-

tribution of liability for defective product damage in court, once it has been admitted – as we have attempted to do quickly in the previous paragraphs – that a defect is also a Cybersecurity deficit in addition to the possible negative impact of the AI system on individuals, groups and society.

The ever-increasing importance of international technical standards, including for AI systems, could, in fact, provide manufacturers who comply with them correctly with what has been defined by the as a 'presumption of conformity' with legal requirements, technically defining the levels of safety expected in order not to be considered defective.

This reduces legal uncertainty and facilitates access to the single market. The prospect appears significant when one considers that regulations such as NIS2 establish that management bodies are personally responsible for approving cyber risk management measures and may incur penalties in the event of non-compliance.

The Cybersecurity regulatory framework can be compared to an emergency braking system in an autonomous car: the law not only prevents the car from leaving the factory with broken brakes (manufacturing defect), but also obliges the manufacturer to intervene remotely if a fault is discovered while the car is already on the motorway (post-market updates). If the manufacturer does not send the repair signal or if the brakes do not meet certified test standards (ISO standards), in the event of an accident, the responsibility lies entirely with the manufacturer of the system, relieving the passenger of the burden of explaining the exact error in the software code. If, then, one also wishes to give weight to non-compliance with the guidelines of ISO 42005:2025, the thresholds of liability rise further and, at present, it would be appropriate to reflect on this in order to determine them more precisely.

5. Towards a model of mandatory pre-market ISO certification for damage prevention

The growing complexity of digital systems and the information asymmetry between manufacturers and consumers make a model based solely on self-assessment of compliance increasingly inadequate. In this perspective, it seems desirable to move towards mandatory pre-market ISO certification schemes, entrusted to independent third-party bodies, with a view to preventing harm. More specifically, the integration of Cybersecurity certification schemes, based on international standards and verified by independent third-party bodies, would offer a solid and uniform 'presumption of conformity' throughout the single market.

The Cybersecurity Act (Reg. EU 2019/881) has already laid the foundations for a European Cybersecurity certification system for ICT products and services; the integration of these schemes into the CRA and PLD framework would strengthen consumer protection and realign civil liability with a genuinely preventive approach. Technical certification would thus take on a function similar to compulsory testing in road traffic law: not a mere formal requirement, but a substantive condition for the legitimacy of placing the product on the market.

In conclusion, the new European law on Cybersecurity and producer liability, together with that on AI systems, marks the transition from a reactive to a preventive model, in which technological security becomes a prerequisite for trust, competitiveness and the sustainability of innovation in the digital market.

ENRICO AL MUREDEN*

AI Act and automated driving between current developments and experimentation§

Abstract

This contribution explores the erosion of the traditional mobility model and the subsequent regulatory transformation driven by the proliferation of Advanced Driver Assistance Systems (ADAS) and highly automated vehicles. Adopting a risk-based regulatory approach consistent with the provisions of the AI Act (Regulation (EU) 2024/1689), the paper analyzes the five levels of driving automation and the pursuit of the "Vision Zero" objective to eliminate road-traffic injuries. The analysis highlights the fundamental role of regulatory sandboxes and real-world testing in fostering responsible innovation while safeguarding individuals in a position of technological vulnerability. Furthermore, it addresses the recalibration of civil liability, examining the transition from a driver-focused to a product-focused liability model. From a de iure condendo perspective, the study evaluates the allocation of accident costs for Level 4 and Level 5 vehicles, contrasting the preemption doctrine with the introduction of a Market Enterprise Responsibility (MER) system designed to ensure the socialization of risks and effective victim compensation.

Il contributo esamina il progressivo superamento del modello tradizionale di mobilità e la conseguente trasformazione del quadro regolatorio determinata dalla diffusione dei sistemi avanzati di assistenza alla guida

* Full Professor of Civil Law at Alma Mater Studiorum-Università di Bologna.

§ This contribution develops the presentation at the conference 'Cybersecurity e prodotti: sfide e responsabilità nelle società interconnesse', 29 November 2024, Italian Supreme Court.

(Advanced Driver Assistance Systems – ADAS) e dei veicoli ad elevato livello di automazione. Muovendo da un approccio regolatorio fondato sul rischio, in linea con le previsioni dell'AI Act (Regolamento (UE) 2024/1689), lo studio analizza i cinque livelli di automazione della guida e la tensione teleologica verso il conseguimento dell'obiettivo "Vision Zero", volto all'azzeramento delle vittime e dei feriti gravi da sinistri stradali. L'analisi mette in luce il ruolo centrale dei regulatory sandboxes e della sperimentazione in condizioni reali quali strumenti funzionali alla promozione di un'innovazione responsabile, capace di coniugare lo sviluppo tecnologico con la tutela dei soggetti esposti a forme di vulnerabilità tecnologica. Inoltre, il contributo affronta il tema della ricalibratura dei modelli di responsabilità civile, soffermandosi sul passaggio da un paradigma incentrato sulla figura del conducente a un as setto maggiormente orientato alla responsabilità da prodotto.In una prospettiva de iure condendo, lo studio valuta i criteri di allocazione dei costi del sinistro per i veicoli di Livello 4 e Livello 5, ponendo a confronto la dottrina della preemption con l'introduzione di un sistema di Market Enterprise Responsibility (MER), concepito al fine di garantire la socializzazione del rischio e un'effettiva tutela risarcitoria delle vittime.

1. Premises

The need to regulate the multiple forms of experimentation currently unfolding in the automotive sector is particularly pressing at a historical juncture in which the traditional mobility model – based on the human driver – is undergoing constant erosion. This erosion is driven by the rapid proliferation of assisted-driving technologies, implemented through Advanced Driver Assistance Systems (ADAS), which constitute the prelude to a future increasingly oriented towards connected, highly automated and, ultimately, autonomous driving[1].

From this perspective, a comprehensive revision of the legal framework governing road-traffic safety, vehicle quality and safety, infrastructure design and the involvement of all road users – including pedestrians, cyclists and users of micro-mobility devices – has become essential. Although these categories are not subject to the licensing procedures required for driving, they nonetheless represent a fundamental component of contemporary mobility and its overall safety structure.

Within this complex scenario, the progressive emergence of technologies relying, inter alia, on Artificial Intelligence systems appears to offer significant benefits in terms of enhanced road-traffic safety. At the same time,

[1] G. Calabresi, E. Al Mureden, *Driverless cars. Intelligenza artificiale futuro della mobilità,* Il Mulino, Bologna, 2021, p. 95.

however, such technologies may pose risks that require careful assessment, both in view of the future consolidation of highly automated mobility and in the current phase, which remains characterized by the coexistence of a vast majority of traditional vehicles alongside next-generation ones.

In this context, several major interventions by the EU legislator come into play, notably the AI Act, Regulation (EU) 2024/1689 of the European Parliament and of the Council[2], and Directive (EU) 2024/2853 of the European Parliament and of the Council.

The new regulatory framework – together with the constant, daily transformations brought about by technological evolution in the automotive sector – requires a profound reinterpretation of the current system. Moreover, from a *de iure condendo* perspective, it calls for a methodological approach that departs from traditional interpretative schemes and necessarily integrates legal and technical viewpoints, adopting a proactive stance. Such an approach demands the continuous monitoring of factual developments, technical innovations and regulatory updates-developments which compel every legal actor, and the legislator foremost among them, to engage in constant revision of rules increasingly exposed to rapid obsolescence.

2. The five levels of driving automation between technical feasibility and regulatory transformation

As policymakers grapple with the task of regulating the current transitional phase – spanning partially automated driving, on-road experimentation of highly automated vehicles, and the resolution of complex issues associated with the future deployment of fully autonomous vehicles – the classification of the five levels of automation developed by the Society of Automotive Engineers International (SAE)[3] plays a pivotal role.

[2] G. Alpa, *Quale modello normativo europeo per l'intelligenza artificiale?,* in «Contratto e impresa», IV (2021), p. 1003; G. Finocchiaro, *Intelligenza artificiale e responsabilità,* in «Contratto e impresa», II (2020), p. 713; V. Di Gregorio, *Intelligenza artificiale e responsabilità civile: quale paradigma per le nuove tecnologie?,* in «Danno e resp.», (2022), p. 51; G. Di Rosa, *Quali regole per i sistemi automatizzati "intelligenti"?,* in «Riv. dir. civ.», V (2021), p. 823; M. Franzoni, *La responsabilità civile una lunga storia ancora da scrivere,* in «Contratto e impresa» IV (2021), in part. p. 1121; Id., *La responsabilità civile fra sostenibilità e controllo delle attività umane,* in «Danno e resp.», IV (2022), p. 5 ff., in part. p. 11.

[3] SAE International (www.sae.org) is a standards organization in the aerospace, automotive, and vehicle industries that develops technical standards to support the manufacturing of products and the performance of activities.

This taxonomy identifies:

– Level 0, designating traditional vehicles in which driving is entirely performed by the human driver;

– Level 1, characterized by driver-assistance systems (such as Stability Control, Cruise Control, Lane-Keeping Technology), now ubiquitous in most modern vehicles;

– Level 2, covering vehicles equipped with automated systems capable of controlling both steering and acceleration/deceleration under predefined conditions, but which nonetheless require continuous human supervision.

The subsequent levels – although fully achievable from a technical standpoint – concern vehicles presently undergoing experimentation across jurisdictions that exhibit markedly different regulatory frameworks. Testing occurs both in controlled environments and on public roads.

Vehicles at:

– Level 3, rely on systems capable of managing all aspects of driving within mapped environments, yet still require a human driver on board to monitor operations and intervene whenever real-world conditions diverge from the pre-mapped scenario, thus necessitating actions that could not be pre-programmed ex ante.

– Level 4, include vehicles in which the presence of a human driver is unnecessary under all operating conditions; the availability of human-driving controls serves only the residual purpose of allowing manual driving in environments where automated operation is impossible or where the user chooses to drive.

– Level 5, vehicles, finally, depart entirely from the conventional concept of a motor vehicle: they perform all driving functions autonomously, exclude any human involvement, and are thus devoid of driving controls such as steering wheels or pedals.

In this evolving landscape, technological innovation is profoundly reshaping vehicular circulation, giving rise to an inevitable period of gradual transition in which traditional vehicles will progressively coexist with increasingly automated ones. A new taxonomy of accidents will emerge: collisions involving only traditional vehicles; collisions between traditional and automated vehicles; accidents involving solely automated vehicles; accidents in which automated vehicles cause harm to pedestrians or cyclists; and, finally, single-vehicle crashes in which an automated vehicle causes damage only to its occupants (the so-called one-car collision)[4].

The task of recalibrating the relationship between driver liability and producer liability – given the spread of driver-assistance and automated-dri-

[4] E. Al Mureden, *Diritto dell'automotive*, Il Mulino, Bologna, 2024, p. 124.

ving technologies – may be addressed from three distinct perspectives, corresponding to the phases marking the transition from a driver-focused to a product-focused liability model.

1. First, long-standing questions arise concerning the impact on civil-liability systems of the widespread adoption of ADAS technologies corresponding to Level 1 and basic Level 2 features.

2. Second, further issues emerge regarding the liability regime applicable to accidents caused by automated vehicles equipped with functionalities falling under Conditional Automated Driving Functionalities (advanced Level 2 and Level 3), within the contexts and limits in which their circulation is permitted.

3. Finally, in a de iure condendo perspective, it appears indispensable to establish guiding principles for the creation of a new body of rules capable of governing the multifaceted aspects associated with the circulation of Level 4 and Level 5 vehicles, in which human intervention will become increasingly marginal, if not effectively residual (Fully Automated Systems)[5].

3. The "Vision Zero" objective and the containment of human error between Advanced Driver Assistance Systems (ADAS) and driverless cars

The rapid evolution of technologies that, in the near future, will enable the widespread deployment first of assisted-driving vehicles and subsequently of driverless cars will lead to a transformation of mobility in which values traditionally regarded as mutually incompatible may finally be reconciled. The emergence of a model of circulation based on driverless, electrified and connected vehicles will indeed make it possible to pursue the goals of emissions reduction, the reorganization of both urban and extra-urban spaces, and – also through a widespread redesign of road networks – the substantial elimination of road-traffic injuries[6].

[5] E. Al Mureden, voce *Responsabilità da circolazione di veicoli,* in *Enciclopedia del diritto,* i Tematici, Responsabilità civile, diretto da C. Scognamiglio, Giuffrè, Milano, 2024, pp. 1052-1076.; A. Albanese, *La responsabilità civile per i danni da circolazione di veicoli ad elevata automazione,* in «Eur. e dir. priv.», IV (2019) p. 995; M.C. Gaeta, *Liability rules and self-driving cars. The evolution of tort law in the light of new technologies,* Editoriale Scientifica, Naples, 2023; S. Pellegatta, *Profili civilistici della «mobilità intelligente»,* Giappichelli, Turin, 2022.

[6] For an accurate description of the policies set out by the EU to achieve the 'zero fatalities' target by 2050, see: https://www.europarl.europa.eu/news/it/press-room/20210930IPR13926/zero-morti-sulle-strade-ue-entro-il-2050

In such a scenario, the opportunity to overcome the constraints inherent in the traditional conception of driving, and to create the conditions for a rapid, reliable and safe transition towards the widespread use of automated-driving vehicles, has been highlighted by the European Commission within GEAR 2030[7], which set out a Roadmap for Automated Driving[8]. Similar considerations appear in a subsequent report commissioned by the European Parliament, which emphasizes the need for a systematic revision of the existing regulatory frameworks and the introduction of rules tailored to the specific risks associated with the circulation of such vehicles – risks whose proper management is viewed as a driver of economic growth. This growth, however, necessarily requires the involvement of a broad and interconnected network of sectors linked to road-traffic circulation, as reflected in the Platform for the Deployment of Cooperative Intelligent Transport Systems in the European Union (C-ITS)[9].

The path undertaken by the European Union is thus directed towards the identification of uniform rules capable of promoting experimentation and, subsequently, the dissemination of highly automated vehicles within a harmonized legal framework. Such harmonization seeks to minimize divergences among national legal systems in all areas affecting vehicular circulation, thereby ensuring the reliability of traffic scenarios in which decisions traditionally taken by human drivers will increasingly result from interaction between humans and automated mechanisms – or may, in some cases, be entirely delegated to the latter. In this respect – particularly in light of the profound regulatory changes introduced by the AI Act (Regulation (EU) 2024/1689) and Directive (EU) 2024/2853 of the European Parliament and of the Council, which repeals Directive 85/374/EEC[10] – it becomes indispensable to undertake a systematic revision of

[7] The European Commission established the High Level Group on the Competitiveness and Sustainable Growth of the Automotive Industry in the European Union (GEAR 2030) in October 2015 in order to bring together the various interests underlying road transport. The group includes representatives of the Member States and stakeholders reflecting the concerns of industry, consumers, environmental protection, and road safety, with the aim of outlining the policies to be adopted in view of the 2030 time horizon.

[8] See the document Prioritising the Safety Potential of Automated Driving in Europe, published by the European Transport Safety Council (ETSC) https://etsc.eu/.

[9] In this regard, reference is made to the document https://ec.europa.eu/transport/sites/transport/files/themes/its/doc/c-its-platform-finalreport-january-2016.pdf, which summarises the needs that emerged from the consultation of a diverse group of stakeholders.

[10] On the burden of proof in product-liability cases, see A. Fusaro, *Responsabilità del produttore: la difficile prova del difetto,* in «Nuova giur. civ. comm.», VI (2017), p. 896; Ead., *I prodotti difettosi e pericolosi: le responsabilità*, in La responsabilità del produttore, in G. Alpa (ed.), Giuffrè, Milan, 2019, p. 361, especially p. 383.

the rules governing driver licensing, road-infrastructure technical standards (Road Worthiness Legislation, as regulated by Directive 2014/45/EU), motor-insurance requirements (Directive 2009/103/EC), and the framework on motor-vehicle type-approval (notably Regulation (EU) 2018/858, which repealed Directive 2007/46/EC).

Within this evolving landscape, Regulation (EU) 2019/2144 – stressing the need to enhance the use of new technologies in order to «further improve vehicle safety performance within an integrated approach to road safety and with a view to better protecting vulnerable road users» (Recital 3) – has mandated that vehicles newly registered from July 2022 onwards be equipped with a series of devices intended to increase both active and passive safety (Advanced Driver Assistance Systems, ADAS)[11], as well as to monitor driving behavior through the compulsory installation of an Event Data Recorder (EDR) (Article 6)[12].

Taken together, these systems appear capable of triggering a virtuous circle – an institutional "nudge" – that, by influencing the behavior of drivers, pedestrians and even manufacturers, will accelerate the transition towards mobility that is automated, connected, sustainable and more respectful of personal safety. As will be discussed, however, the actual impact of such systems on overall road-traffic safety can only be assessed accurately after a significant monitoring phase, allowing for observation of both the benefits and the risks linked to the introduction of new driver-assistance technologies in real-world environments.

4. Autonomous driving between regulatory sandboxes and real-world testing

Even more delicate issues arise in connection with technologies still in the experimental phase which – although fully functional in laboratory settings and in dedicated testing environments – have not yet been placed on the market.

[11] The issue, already examined in the European context as early as 2005 (K. Van Wees, K. Brookhuis, *Product liability for ADAS; legal and human factors perspectives*, in «EJTIR», 5, IV (2005), p. 357, has in recent times gained increasing relevance in the United States (see K.S. Abraham, R.L. Rabin, *Automated Vehicles and Manufacturer Responsibility for Accidents: A New Legal Regime for a New Era,* in «Virginia Law Review», (2019), Forthcoming; Stanford Public Law Working Paper, SSRN: https://ssrn.com/abstract=3159525, p. 8). In that system, particular emphasis is placed on the need to address a significant transitional period during which the continued prevalence of vehicles with traditional driving systems will require the current rules governing civil liability to remain unchanged. According to the authors, these rules should undergo a thorough revision only once automated vehicles account for 25% of the circulating fleet.

[12] E. Al Mureden, *Diritto dell'automotive*, cit. p. 124.

In this respect, particular significance should be attributed to the emphasis placed by the AI Act[13] on supporting innovation and promoting research and development activities[14].

The AI Act limits the application of its substantive provisions to the phase in which an Artificial Intelligence system is placed on the market. In the preceding development phase, the rules governing regulatory sandboxes and real-world testing acquire fundamental relevance. These rules safeguard an anthropocentric model for the development of AI systems, which requires regulatory oversight and a safe and controlled environment for experimentation, while simultaneously ensuring responsible innovation and the integration of adequate safeguards and risk-mitigation measures[15]. Accordingly, each Member State is required to establish regulatory sandboxes designed to facilitate the development and testing of innovative AI systems under strict regulatory supervision prior to market placement. These sandboxes – whether physical, digital or hybrid, and accommodating both physical and digital products – correspond, in the automotive sector, to simulation centers and, above all, to so-called smart roads. Within such environments, it becomes possible to pursue the objectives of fostering AI innovation by enabling controlled testing during the development and pre-commercialization phases. Furthermore, these environments make it possible to ensure compliance of innovative AI systems with Regulation (EU) 2024/1689 and with the associated harmonized technical standards.

Additional functions of AI regulatory sandboxes include:

– strengthening legal certainty for innovators and improving the ability of competent authorities to monitor and understand the opportunities, emerging risks and impacts of AI use[16];

– facilitating regulatory learning for both authorities and companies, in legal framework[17].

Ultimately, the establishment of AI regulatory sandboxes serves to reduce the legal uncertainty arising in the pre-market phase. Through careful

[13] *See* Recital n. 25.

[14] It is therefore necessary to exclude from its scope the AI systems and models that are specifically developed and deployed solely for the purpose of scientific research and development. It is also necessary to ensure that the Regulation does not otherwise affect scientific research and development activities relating to AI systems or models prior to their being placed on the market or put into service. As regards product-oriented research, testing, and development activities relating to AI systems or models, the provisions of the Regulation should likewise not apply before such systems and models are put into service or placed on the market.

[15] Recital n. 25.

[16] *Ibidem.*

[17] From this perspective, it is particularly significant that Recital 25 emphasizes the importance of providing adequate support for cooperation and the sharing of best practices

supervision within the sandbox, AI systems undergo development, training, testing and validation before market placement, thereby enabling optimal management of any significant risks identified during testing.

The analysis conducted within regulatory sandboxes complements the rules governing «real-world testing of high-risk AI systems outside AI regulatory sandboxes», as provided under Article 60 of Regulation (EU) 2024/1689. Real-world testing may be carried out by providers or prospective providers of high-risk AI systems listed in Annex III only when a series of cumulative conditions are met conditions relating to:

– prior submission and approval of a real-world testing plan;
– registration of the test;
– establishment within the Union or appointment of a Union-established representative;
– safeguards for data transfers to third countries;
– temporal limits on the duration of real-world testing;
– protection of vulnerable groups;
– obligations of information, cooperation and contractual allocation of responsibilities between providers and deployers;
– informed consent of natural persons participating in the tests;
– effective human supervision;
– the ability to override or disregard AI recommendations or decisions.

Finally, Article 73 of the Regulation imposes a duty to report to the national market-surveillance authority any serious incident identified during real-world testing. Upon such an occurrence, the provider or prospective provider must take immediate mitigation measures or, failing that, suspend the testing until mitigation has been achieved. The provider must also establish procedures for the prompt withdrawal of the AI system following suspension or termination of testing and must notify both the interruption and the final results to the relevant national authority.

With respect to civil liability, the provider is liable for any damage caused during real-world testing.

Particular attention must also be paid to informed consent, as governed by Article 61. Consent must be given freely by participants, who must be duly informed – clearly, concisely and in a comprehensible manner – of the

with the authorities involved in the regulatory experimentation space for AI, as well as accelerating access to markets, including the removal of obstacles for SMEs, including start-ups. The recital continues by stating that regulatory experimentation spaces for AI should be widely available throughout the Union, and particular attention should be paid to their accessibility for SMEs, including start-ups.

nature and purpose of the testing; foreseeable inconveniences; conditions and duration; the rights afforded to participants, including the right to refuse participation or to withdraw at any time without justification; and the procedures for requesting that AI-generated recommendations or decisions be ignored or overridden.

5. Automated driving and new frontiers of civil liability

The most complex issues – and those requiring the creation of new legal rules – arise in the scenario in which Level 4 and Level 5 driverless vehicles are placed on the market as products and can operate by fully exploiting their automation capabilities.

As anticipated, full implementation of Level 4 and Level 5 automation is expected, according to the most authoritative studies, to reduce the number of accidents by approximately 90% compared with current levels. Within the remaining share of accidents, some will likely result from product defects,

infrastructure deficiencies or abnormal human behavior. Others, however, will be caused by conforming vehicles – i.e. vehicles operating autonomously according to predefined and legislatively validated safety protocols.

The former category may continue to be governed by existing rules on manufacturer liability, road-infrastructure liability and road-user liability. The latter, however, raises the issue of allocating the costs generated by activities carried out in full compliance with regulatory safety standards – activities which the legislator has deemed reasonably safe[18].

Setting aside ethical concerns, the task is to devise a regulatory framework capable of ensuring both legal certainty and efficient allocation of accident costs arising from the commercialization and circulation of highly automated vehicles. In this respect, various de iure condendo solutions have been proposed in the United States.

[18] Ultimately, it can be stated that the use of partially automated systems does not present legal differentiation compared to situations where driver assistance systems (brake assist, power steering, ABS) are used, as there remains a constant need for monitoring and corrective interventions in case of errors. In this sense, M. Geistfeld, *A Roadmap for Autonomous Vehicles: State Tort Liability, Automobile Insurance, and Federal Safety Regulation*, in «California Law Review», Vol. 105, VI (2017), p. 1636; J. Levy, *No Need to Reinvent the Wheel: Why Existing Liability Law Does Not Need to Be Preemptively Altered to Cope with the Debut of the Driverless Car,* in 9 «J. Bus. Entrepreneurship & L.», (2015), p. 360; K.S. Abraham, R. L. Rabin, op. cit., p. 40, argue that in the near future, the issue of the proliferation of vehicles with level 3 automation will mainly raise problems related to a malfunction of the algorithm responsible for drawing the driver's attention and prompting their intervention. Since the manufacturer's liability could only be configured in this regard, it is possible to conclude that even with

A first, strongly pro-automation approach – aimed at accelerating the replacement of human-driven vehicles with automated ones – relies on a coordinated reading of federal vehicle-safety regulations and state tort-law provisions. Under this view, the preemption doctrine[19] plays a pivotal role: conformity with federal safety standards established by the National Highway Traffic Safety Administration (NHTSA) acts as a limit beyond which no civil liability may be imposed on a manufacturer. Developed under the National Traffic and Motor Vehicle Safety Act of 1966 with regard to traditional vehicles, this doctrine would extend to future federal legislation on highly automated vehicles[20].

Such an interpretation would elevate regulatory compliance to a complete defense, thereby promoting the rapid diffusion of safer technology. Manufacturer liability would thus arise only where the vehicle is non-compliant with federal standards – i.e. where accidents result from product defects, hardware malfunctions or failures to provide adequate warnings enabling reasonably safe use. Under this model, the burden of compensating third-party victims of accidents caused by compliant automated vehicles would continue to fall primarily on the vehicle owner, in their capacity as owner or user. In one-

the spread of level 3 vehicles, there is no need for a radical modification of the rules currently governing civil liability. In the Italian context, the issue of identifying rules capable of responding to the new demands generated by the transition from a human-driven traffic model to an automated driving model is analyzed by U. Ruffolo, *Self-driving car, auto driverless and responsibility, in Intelligenza artificiale e responsabilità*, a cura di Id., Giuffrè, Milan, 2017, p. 39. The Author highlights the possibility of resorting to the provisions of Article 2050 of the Civil Code regarding hazardous activities, emphasizing the specific dangers associated with highly automated driving. Conversely, A. Davola, R. Pardolesi, *Traveling with the robot: towards new horizons of auto liability ("driverless")?,* in «Danno e resp.», V (2017), p. 625, exclude the possibility of applying Article 2050 of the Civil Code, as adhering to this interpretation would «end up characterizing an activity (automated driving) as dangerous, which, if compliant with technical standards, is paradoxically safer than normal 'human' driving». Similar considerations are echoed in Ferrari, *Comparative analysis on civil liability related to the circulation of autonomous vehicles,* op. cit., p. 110, where she observes that, with the exception of the United Kingdom, «*in the states where the experimentation of autonomous vehicles has been authorized by law, the insurance coverage limits have been significantly increased*». Such a choice is observed not so much from the perspective of «recognizing greater danger of autonomous vehicles» but rather as a measure aimed at facilitating the acceptance of experimentation by the public.

[19] About preemption, see T.O. McGarity, *The Preemption War: When Federal Bureaucracies Trump Local Juries,* New. Haven-London, (2008); A.E. Untereiner, *The preemption defense in tort actions: Law Strategy and Practice*, Washington D.C., 2008; J.T. O'Reilly, *Federal Preemption of State and Local Law,* American Bar Association, Chicago, 2006; J.L. Goldsmith, *Statuory Foreign Affairs Preemption*, in «The Supreme Court Review», (2000); C.M. Sharkey, Products Liability Preemption: an Institutional Approch, in «Geo. Wash. L. Rev»., (2008); M.J. Davis, *On Restating Products Liability Preemption*, in «Brooklyn L. Rev.», Vol 74, III (2009).

[20] B.W. Smith, *Automated Driving and Product Liability*, in «Mich. St. L. Rev.» I (2017) available at SSRN: https://ssrn.com/abstract=2923240; M. Geistfeld, op. cit., 1674.

car collisions, where a compliant automated vehicle causes injury solely to its owner-occupant, the preemption doctrine would similarly bar recovery[21].

This approach mirrors the rules governing liability for conforming products in various market sectors. While seemingly advantageous for manufacturers by limiting exposure to defective-product cases only, it risks undermining consumer confidence, amplifying the well-known status quo bias.

A contrasting approach highlights the need to socialize the costs of the "new" accidents arising from automated mobility. This vision supports the introduction of a Market Enterprise Responsibility (MER) system, financed through a fund contributed to by manufacturers and intended to provide compensation for victims of accidents caused by driverless vehicles[22]. Such a mechanism would offer broadly indemnity-based coverage for occupants, pedestrians and other bystanders, and would distribute compensation costs according to the principle of market share liability, whereby each manufacturer contributes proportionally to its market share. The contribution would therefore reflect the risk introduced into society by each manufacturer, measured by the volume of vehicles sold.

The MER model would need to be supplemented by a further parameter: the safety coefficient of each manufacturer's fleet. Manufacturers with vehicles demonstrating higher levels of active and passive safety – i.e. greater ability to prevent accidents or mitigate their consequences – would bear lower

contributions than manufacturers achieving the same market share but with less safe vehicles[23].

If, on the one hand, it is essential to promote technological development – also through legislative and judicial interpretations that may encourage investment in Highly Automated Vehicles – on the other, it is equally necessary to establish adequate mechanisms for compensating those exposed to risks that cannot yet be fully assessed during the ongoing experimentation phase.

For these reasons, scholarly proposals advocating a cautious *de iure condendo* approach and calling for instruments that prevent accident costs – even those caused by conforming vehicles – from falling upon victims, appear particularly persuasive. Compensation may need to be ensured sometimes through civil liability, sometimes through broadly indemnity-based mechanisms, and sometimes through a synergistic combination of both.

In a future dominated by driverless vehicles, a distinction must be drawn between damage caused to third parties and damage suffered by the vehicle owner/user. As regards third-party damage, the current civil-liability fra-

[21] *Ibidem.*

[22] K.S. Abraham, R.L. Rabin, cit., p. 44; M. Geistfeld, cit., p. 1635; A. Albanese, cit., p. 995.

[23] G. Calabresi, E. Al Mureden, cit., p. 95.

mework may simply be adapted by assigning greater relevance to the owner's liability and progressively reducing the significance of the driver's negligence — thus applying a rule analogous to Article 2054(4) of the Italian Civil Code[24]25.

For damage suffered by the owner/user, however, a profound revision of product-liability rules becomes necessary. In this context, abandoning the principle that "there is no liability without defect" – which mirrors the U.S. pre-emption doctrine – in favor of a compensation model such as the Market Enterprise Responsibility would facilitate victim compensation. This is particularly appropriate in the presence of products whose complexity renders the traditional requirement of proving defect increasingly inadequate, and in view of the need to reduce the liability gap – one of the central objectives of the future EU framework on producer liability.

[24] E. Al Mureden, voce *Responsabilità da circolazione di veicoli*, cit., p. 1052.

BIBLIOGRAPHY

Abraham K.S., Rabin R.L., *Automated Vehicles and Manufacturer Responsibility for Accidents: A New Legal Regime for a New Era,* in «Virginia Law Review», Vol. 105, (2019), Forthcoming; Stanford Public Law Working Paper, SSRN: *https://ssrn.com/abstract=3159525*

Albanese A., *La responsabilità civile per i danni da circolazione di veicoli ad elevata automazione*, in «Eur. e dir. priv.», IV (2019), p. 995.

Al Mureden E., *Diritto dell'automotive,* Il Mulino, Bologna, 2024.

Al Mureden E., voce *Responsabilità da circolazione di veicoli*, in Enciclopedia del diritto, i Tematici, Responsabilità civile, diretto da Scognamiglio C., Giuffrè, Milano, 2024, pp. 1052-1076.

Alpa G., *Quale modello normativo europeo per l'intelligenza artificiale?,* in «Contratto e impresa», IV (2021), p. 1003.

Calabresi G., Al Mureden E., *Driverless cars. Intelligenza artificiale futuro della mobilità*, Il Mulino, Bologna, 2021.

Davis M.J., *On Restating Products Liability Preemption*, in «Brooklyn L. Rev.», Vol. 74, III (2009).

Davola L., Pardolesi R., *Traveling with the robot: towards new horizons of auto liability ("driverless")?*, in «Danno e resp.», V (2017), p. 625.

Di Gregori L., *Intelligenza artificiale e responsabilità civile: quale paradigma per le nuove tecnologie?,* in «Danno e resp.», (2022), p. 51.

Di Rosa G., *Quali regole per i sistemi automatizzati "intelligenti"?,* in «Riv. dir. civ.» V (2021), p. 823.

Dinh V.D., *Reassessing the law of Preemption*, in «Geo L.J.», Vol. 88, VII (2000).

Finocchiaro G., *Intelligenza artificiale e responsabilità*, in «Contratto e impresa», II (2020), p. 713.

Franzoni M., *La responsabilità civile una lunga storia ancora da scrivere*, in «Contratto e impresa» IV (2021), in part. p. 1121;

Franzoni M., *La responsabilità civile fra sostenibilità e controllo delle attività umane*, in «Danno e resp.» IV (2022), p. 5.

Fusaro A., *Responsabilità del produttore: la difficile prova del difetto*, in «Nuova giur. civ. comm.», VI (2017), p. 896.

Fusaro A., *I prodotti difettosi e pericolosi: le responsabilità*, in *La responsabilità del produttore*, in G. Alpa (ed.), Giuffrè, Milan, 2019, p. 361, especially p. 383.

Gaeta M.C., *Liability rules and self-driving cars. The evolution of tort law in the light of new technologies*, Editoriale Scientifica, Naples, 2023.

Geistfeld M.A., *A Roadmap for Autonomous Vehicles: State Tort Liability, Automobile Insurance, and Federal Safety Regulation*, in «California Law Review», Vol. 105, VI (2017), p. 1636.

Goldsmith J.L., *Statuory Foreign Affairs Preemption*, in «The Supreme Court Review», (2000).

Levy J., *No Need to Reinvent the Wheel: Why Existing Liability Law Does Not Need to Be Preemptively Altered to Cope with the Debut of the Driverless Car*, in «J. Bus. Entrepreneurship & L.», Vol. 9, II (2016), p. 360.

McGarity T. O., *The Preemption War: When Federal Bureaucracies Trump Local Juries*, New. Haven-London, 2008.

J.T. O'Reilly, *Federal Preemption of State and Local Law*, American Bar Association, Chicago, 2006.

Pellegatta S., *Profili civilistici della "mobilità intelligente",* Giappichelli, Turin, 2022.

Ruffolo U., *Self-driving car, auto driverless and responsibility*, in *Intelligenza artificiale e responsabilità*, a cura di Id., Giuffrè, Milan, 2017, p. 39.

Sharkey C. M., Products Liability Preemption: an Institutional Approch, in «Geo. Wash. L. Rev.», Vol. 76, (2008).

Smith B.W., *Automated Driving and Product Liability*, in «Mich. St. L. Rev.», I (2017) available at SSRN: https://ssrn.com/abstract=2923240;

A.E. Untereiner, *The preemption defense in tort actions: Law Strategy and Practice*, Washington D.C., 2008.

K. Van Wees, K. Brookhuis, *Product liability for ADAS; legal and human factors perspectives*, in «EJTIR», IV (2005), p. 357.

ANTONINA ASTONE*

Cyber Security and Healthcare Facilities§

Abstract

This essay examines current legislation on personal data protection, with particular reference to health data, and the risks arising from a breach. The analysis focuses on the central role of Cybersecurity within the European data strategy, including regulatory developments resulting from the enactment of the NIS 2 Directive and the AI Act, taking into account the alternative offered by blockchain in terms of data security and robustness.

Il presente saggio esamina l'attuale legislazione in materia di protezione dei dati personali, con particolare riferimento ai dati sanitari, e i rischi derivanti da una loro violazione. L'analisi si concentra sul ruolo centrale della Cybersicurezza all'interno della strategia europea dei dati, includendo gli sviluppi normativi derivanti dall'adozione della Direttiva NIS 2 e dell'AI Act, tenendo conto dell'alternativa offerta dalla tecnologia blockchain in termini di sicurezza e robustezza dei dati.

1. Introduction

Cybersecurity in healthcare plays a crucial role as there is a directly proportional relationship between the security of computer systems[1], gene-

* Professor of Private Law at the University of Messina.

§ This contribution develops the presentation at the conference 'Cybersecurity e prodotti: sfide e responsabilità nelle società interconnesse', 29 November 2024, Italian Supreme Court.

[1] According to data provided by the National Cybersecurity Agency www.acn.gov.it, in Italy, in the healthcare sector in 2023 there was an increase in cases of violation of IT systems with 60% of them being particularly serious. Cybercrime represented the dominant typology (93%).

rally understood and used within healthcare facilities, and the health if not the very life of patients. This is demonstrated by the case of a woman who died in 2020 following a cyber attack that had paralyzed the hospital in Düsseldorf, Germany. The delay in treatment, due to the transfer to another hospital, due to an computer system block, proved fatal.

The resilience of the healthcare sector against cybercrime attacks is a strategic objective of the European legislator both to protect patient data and to ensure the continuity of essential services. The reason why the healthcare system is attractive to cybercriminals lies, evidently, in the very nature of the data processed, called, not by chance, in the previous formulation of the Italian privacy code «sensitive»[2], as they concern the most intimate sphere of the person.

Undue knowledge of these particular data may determine, for the individual, prejudices linked, for example, to exclusion from insurance coverage or access to credit or, again, cause discrimination in the workplace. It follows that their treatment must take place by trying to exclude or minimize as much as possible the risk of their theft or undue access or improper use which produces, in fact, a strong negative impact on fundamental rights.

At European level, the reference regulations are all those relating to the «data strategy»[3] which starts from the GDPR Regulation 679/2016, General Data Protection Regulation, unfolds through the regulation on the processing of non-personal data[4], and the one on data governance up to the Data Act[5], and the AI Act[6], as well as the regulation on Cybersecurity[7].

The common thread of the interventions mentioned is to balance the protection of fundamental rights with the need for the processing, use and cir-

[2] Health data now fall within the special categories of personal data referred to in Article 9, paragraph 1, of EU Reg. 2016/679.

[3] Brussels, 19.2.2020 COM (2020) 66 final Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, A European Strategy for Data.

[4] Reg. EU n. 1807/2018 regarding « a framework for the free flow of non-personal data in the European Union».

[5] EU Regulation 2022/868 of the European Parliament and of the Council of 30 may 2022 on European data governance and amending Regulation (EU) 2018/1724 Data governance act; EU Reg. 2023/2854 concerning harmonised rules on fair access to and use of data and amending Regulation (EU) 2017/2394 and Directive (EU) 2020/1828 s.c. Data act.

[6] Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 june 2024 laying down harmonised rules on Artificial Intelligence. In doctrine P. VOIGT, N. HULLEN, The Eu AI act, Berlino-Heidelberg, Springer, 2024. U. RUFFOLO (a cura di), AI Act. *La regolamentazione europea dell'intelligenza artificiale*, Roma, Luiss University press, 2025.

[7] Regulation (EU) 2019/881 of the European Parliament and of the Council of 17 april 2019 on ENISA (the European Union Agency for Cybersecurity) and on information and communications technology Cybersecurity certification, and repealing Regulation (EU) No 526/2013.

culation of data for the functioning of the digital ecosystem. The starting point is therefore the reference to the «social function»[8] which already in recital 4 of the GDPR, is assigned to the right to the protection of personal data which cannot, therefore, be considered an « absolute prerogative» (cons. 4), as it involves a super-individual evaluation of the processing.

It seems superfluous to recall, in fact, the crucial role played by data in general in the current economic contingency also for the development of goods and services in the private and public sectors. In particular, the digitalization of health data has been a key resource for implementing the efficiency of the health system, especially in light of the increasingly frequent use of Artificial Intelligence. The positive effects, also in terms of predictive medicine, connected to the use of increasingly smart medical devices, as well as the collection of data in the electronic file, in terms of greater efficiency in the provision of services, are therefore evident.

Generally, health data are processed using advanced pseudonymization and anonymization techniques, in order to guarantee the privacy of the owners in accordance with the security of the data, which is a fundamental prerequisite for proceeding with their processing.

2. The social function performed by data and the measures for their safe treatment

Examining the reference legislation, it emerges that in the GDPR, personal data breach referred to in art. 4, means «*a breach of security leading to the accidental or unlawful destruction, loss, alteration, unauthorized disclosure of, or access to, personal data transmitted, stored or otherwise processed*». In 2018, in particular, the Working Party 29 drafted guidelines on Data Breach, distinguishing three subcategories: the «Confidentiality Breach», in the event of accidental or abusive access to personal data; «Availability Breach», when there is an accidental or unauthorized loss or destruction of personal data; «Integrity Breach», when there is an accidental or unauthorized alteration of personal data.

Generally in healthcare facilities[9], such as hospitals, there is a Data Protection Officer and a manager who can be internal such as, for example, the

[8] R.A. Epstein, *Takings-Private property and the power of eminent domain*, Harvard University Press, Cambridge, 1985, p. 166 ff., which investigates the relationship between public and private interests underlying the right to property, in order to demonstrate their possible convergence within the law itself.

[9] W. Gordon, A. Fairhall, A. Landman, *Threats to Information Security-Public Health Implications.* in «New England Journal of Medicine», (2017), pp.707; C.S. Kruse, B. Frederick, D. Montales, *Cybersecurity in healthcare: A systematic review of modern threats and trends*, in «Technology and Health Care», (2017), pp.12; E. Perakslis, *Cybersecurity in health care*, in «New England Jour-

Director of the facility or external as the security in data processing can also be entrusted to an external subject who can be bound by a contract with the data controller pursuant to art. 28 and who assumes the related responsibility. Other key figures are then the subjects authorized to process data who can be subordinate workers such as doctors or paramedical staff. Among the most important obligations of the data controller, is that of evaluating the potential risks for the rights of the interested parties and implementing adequate technical and organizational measures to address them. Furthermore, pursuant to art. 32 of the GDPR, the data controller, « must periodically update the security measures that protect the processing of personal data». To this end, pursuant to art. 35 of the GDPR, it is necessary to carry out a personal data protection impact assessment (DPIA), a fundamental requirement in cases of more risky processing[10].

In the context of balancing data protection and the need for their use and circulation, the Data Governance Act also the Data act always start from the perspective of the social function performed by data, both personal and non-personal, aiming to break the data oligopoly situation that characterizes the digital market, including that of healthcare.

The introduction of the obligation to make data available also to third parties authorized by the user, provided for in the Data governance act, will inevitably lead to the need to significantly increase the IT security requirements. It is interesting to underline how the Data act identifies among the public emergencies, pursuant to recital 57, serious IT security incidents.

The healthcare sector will tend, almost entirely, to be included in the high risk levels which, according to the AI act, concern all those Artificial Intelligence systems that can generate a « significant harmful impact on health, safety and fundamental rights». In addition to underlining the need for AI systems to be resilient to attempts by third parties with malicious intent who, by exploiting the system's vulnerabilities, aim to alter its use, behavior, performance or compromise its security properties, the AI Act is concerned, for

nal of Medicine», (2014), pp.395. Generally S.C. Ayala, *Cybersecurity for Hospitals and Healthcare Facilities,* New York, Apress, 2016; R.J. Campbell, *Cybersecurity Vulnerabilities and Considerations* in U.S. Healthcare Facilities: A Scoping Review, www.digitalcommons.com, (2024).

[10] Once a Data Breach has occurred, according to the GDPR, the obligation to notify the data breach to the Privacy Authority and, where possible, to the data owner is triggered. The subject whose data has been breached may also take action to obtain compensation for the damages suffered. Notification of any data breach shall be made without undue delay and, where feasible, not later than 72 hours after becoming aware of the breach. In some cases, where the personal data breach is likely to result in a high risk to the rights and freedoms of natural persons (individually or, if this would require «disproportionate effort», collectively) communicate the breach to the persons whose personal data have been affected by the breach, pursuant to art. 34 GDPR, also in order to allow them to take appropriate precautions to mitigate possible adverse effects resulting from the data breach.

example, that the data used for training are not altered maliciously or through negligence, in such a way as to generate incorrect outputs with harmful effects in the healthcare field. Technical solutions, aimed at addressing AI-specific vulnerabilities, include, where appropriate, measures to prevent and control cyber attacks that seek to manipulate the training dataset (data poisoning) or the trained models (adversarial attacks), or model flaws.

The so-called deployers are considered equivalent to them, which, pursuant to Art. 3, par. 3 of the AI Act, are «a natural or legal person, a public or private body, which uses an AI system under its own responsibility». In the healthcare sector, the deployer will be the healthcare facility or the doctor himself, depending on who assumes responsibility for the use of the AI system and the healthcare service provided through it. If the healthcare contract is stipulated by the healthcare facility, the latter becomes the deployer, while the doctor is considered part of the staff responsible for using the system. If, on the other hand, the doctor stipulates the contract directly with the patient, he himself will be the deployer, assuming both the role of user and that of system manager.

In order to develop trustworthy AI systems, the AI Act introduces an additional impact assessment focused on fundamental rights, the Fundamental Impact Assessment, aimed at assessing the effects that a high-risk AI system may have on the fundamental rights of individuals and identifying the measures to be taken when such risks materialize. The scope of the so-called FRIA seems broader than the required data protection impact assessment (DPIA) under art. 35 of the GDPR, as its greater specificity than the DPIA would not allow it to cover all the aspects that the FRIA needs to focus on.

3. The liability of administrators foreseen by the NIS 2 Directive

The NIS 2, Network and Information Security, Directive completes the European data strategy[11]. This directive issued with the aim of strengthening the protection of digital infrastructures from cyber threats, constitutes a crucial challenge for the European Union and completes the European data strate-

[11] Directive (EU) 2022/2555 of the European Parliament and of the Council of 14 december 2022 on measures for a high common level of Cybersecurity across the Union. Italy implemented the NIS 2 Directive with Legislative Decree no. 138 of 4 September 2024. Shortly before, Law no. 90 of 28 June 2024 was introduced, containing "Provisions regarding the strengthening of national Cybersecurity and the fight against cybercrime". In italian doctrine, generally, G.l. Berruti, G. Corasaniti, P. Perri (eds.), *Nis 2 ed evoluzione della cybersicurezza nazionale. Guida alle norme e alle determinazioni dell'Agenzia per la Cybersicurezza Nazionale*, Maggioli editore, Santarcangelo di Romagna 2025; specifically F.P. Micozzi, *Sicurezza informatica. Obblighi e responsabilità dopo il recepimento della NIS2 e la L. 90/2024*, Wolter Kluwer, Milan, 2024, pp.95.

gy. Healthcare facilities are subject to specific obligations in terms of IT security, risk and incident management.

The healthcare sector is included among those with high criticality and in addition to strengthening the incident notification system, the Nis 2 directive makes the top management of companies and institutions responsible. The most significant innovation introduced is provided for in art. 20 with the provision of responsibility for the failure to comply with safety regulations on the part of company management bodies. This is a significant change from the past: the management of Cybersecurity is no longer a task relegated exclusively to the information systems function, but becomes a direct responsibility of the corporate management body, such as the board of directors of a hospital or the general manager, as well as the administrative and health directors of a hospital facility who would be in addition to that of the office managers.

In healthcare facilities, there are already units dedicated to ensuring the Management Control of Information Systems, but starting from the NIS 2 directive, security is an issue that concerns the entire organization and must be treated as such.

This approach requires a shared commitment and careful management at all levels of the company, to guarantee information security and the resilience of systems in the face of cyber threats[12].

4. The possible alternative offered by blockchain

The rules mentioned so far will inevitably have to be integrated with the upcoming creation of a «single market» for health data, which represents another central objective of the European Strategy for Data[13].

The European Health Data Space intends to broaden the possibilities of access to such data whenever their reuse could have benefits in the health sector.

[12] In par. 1 of art. 20 it is foreseen that national law remains unaffected with regard to the liability rules applicable to public institutions, as well as the liability of public employees and elected or appointed officials. If the fraudulent action of the third party is not an express cause for exemption from liability, neither is the absence of damages for the interested parties following the unlawful intrusion. Therefore, the responsible directors may be subject to administrative pecuniary sanctions, the amount of which must be determined taking into due account the elements provided for in art. 83, par. 2, GDPR. In Italy, the NIS 2 directive was implemented with the Legislative decree of 4 September 2024, n. 138 and security incidents and breaches must be reported to the Computer Security Incident Response Team established at the National Cybersecurity Agency (ACN), without undue delay.

[13] A fundamental piece is the Regulation on the European Health Data Space, the so-called Ehds, whose publication is awaited: in doctrine, W. Li, P. Quinn, *The European Health Data Space: An expanded right to data portability?* in «Computer law & Security Review », vo.52, (2024).

The benefits of its implementation will primarily be to enable faster and easier access to electronic health data, whether people are in their own country or in another Member State. It will also ensure greater control over how this data is used.

The researchers to access specific types of secure and anonymized health data to guide scientific research, develop better treatments and improve patient care. Furthermore, the interoperability of data will enable will mitigate the difficulties related to the different levels of digitalization of health data, which vary from one Member State to another and make it more difficult to share data across borders. All electronic health record systems will have to comply with the provisions of a European electronic health record exchange format. Through a secure platform, the bodies responsible for access to data, designated by the EU Member States, will provide the data to users such as scientific researchers. With regard to the security measures of health data, the Ehds Regulation provides that these users have access for secondary use limited to what is necessary for the purpose of the requested processing, in line with the authorization issued by the responsible body.

But the fact remains that the platforms and, more generally, centralized databases are often the object of «attacks» by cybercriminals, capable of finding increasingly sophisticated lockpicks to violate Cybersecurity protocols.

The GDPR has set itself the goal of giving people back control over their personal data, which, however, is entrusted to other parties, the data controllers, who are responsible for observing a series of rules and procedures for the safe processing of data in light of the various regulations, including the NIS 2 directive.

The risks would be lower if, instead, this data were to pass through blockchain[14]: the decentralization brought by blockchain could change the game of health security. For example, if a hacker were to violate a blockchain, he would be able to access at most a few blocks of data, which are not even usable without the complete context of the chain. Furthermore, thanks to the peer-to-peer sharing of blockchain technology, it is impossible to alter a single block without altering the entire network. Therefore, it is the blockchain itself that provides a community of users capable of raising the alarm when the security of the network has been compromised. This community would therefore be limited to medical personnel. The fact that a medical record on blockchain is a guarantee of integrity, confidentiality and non-alterability also facilitates the investigation phase in the event of a possible judgment of liability. The patients could remain the sole owners of their health data if they were certified as an

[14] S. Nakamoto, *Bitcoin: A peer-to peer electronic cash system*, wwwbitcoin.com, (2008); P. Ruan, T.T.A. Dinh, D. Loghin, M. Zhang, G. Chen, *Blockchains*, Springer, Berlin-Heidelberg, 2022.

NFT, Not Fungible Token, a unique digital certificate, registered in a blockchain, used to certify the ownership of an asset such as, for example, a work of the mind. The introducti could guarantee patients full control over their medical information. It would also prevent such data from being transferred to third parties, without full awareness of the possible consequences, in this sense, by the patients themselves when giving their consent. The NFT, in fact, makes the patient the sole owner of their health data, the only one who can also decide to sell it for profit[15].

[15] M. Finck, *Blockchains and data protection in European union*, in «European Data Protection Law Review», vo. 4, (2018), p. 19.

BIBLIOGRAPHY

Ayala C., *Cybersecurity for Hospitals and Healthcare Facilities,* Apress, New York, 2016.

Berruti G.I., Corasaniti G., Perri P. (eds.), *Nis 2 ed evoluzione della cybersicurezza nazionale. Guida alle norme e alle determinazioni dell'Agenzia per la Cybersicurezza Nazionale,* Santarcangelo di Romagna, Maggioli, 2025.

Campbell R.J., *Cybersecurity Vulnerabilities and Considerations in U.S. Healthcare Facilities: A Scoping Review*, *www.digitalcommons.com,* 2024.

Epstein R.A., *Takings-Private property and the power of eminent domain*, Harvard university press, Cambridge, 1985, p. 166 ff.

Finck M., *Blockchains and data protection in European union,* in «European Data Protection Law Review», (2018), pp. 1-32.

Gordon W., Fairhall A., Landman A., *Threats to Information Security-Public Health Implications*. in «New England Journal of Medicine», (2017), 707-709.

Kruse C.S., Frederick B., Montales D., *Cybersecurity in healthcare: A systematic review of modern threats and trends*, in «Technology and Health Care», (2017), pp. 1-13.

Li W., Quinn P., *The European Health Data Space: An expanded right to data portability?* in «Computer law & Security Review», (2024).

Micozzi F.P., *Sicurezza informatica. Obblighi e responsabilità dopo il recepimento della NIS2 e la L. 90/2024*, Wolter Kluwer, Milan, 2024.

Nakamoto S., *Bitcoin: A peer-to peer electronic cash system,* wwwbitcoin.com, 2008.

Perakslis E., *Cybersecurity in health care*, in «New England Journal of Medicine», vo.371, (2014), pp. 395-397.

Ruan P., Dinh T.T.A., Loghin D., Zhang M., Chen G., *Blockchains*, Springer, Berlin-Heidelberg, 2022.

Ruffolo U. (ed.), AI Act.*La regolamentazione europea dell'intelligenza artificiale*, Luiss University press, Rome, 2025.

Voigt P., Hullen N., *The Eu AI act,* Springer, Berlin-Heidelberg, 2024.

IRENE COPPOLA*[1], CRISTINA MINGO*[2]

Cybersecurity and the protection of vulnerable groups in the digital age: challenges and prevention strategies[§]

Abstract

This contribution offers a brief reflection on the protection of minors and vulnerable individuals who are exposed to the risks of the digital jungle, such as cyberbullying and catfishing. The analysisproposes a holistic approach based on threepillars: 1. a regulatory framework for the implementation of age-verification systems and the enforcement of the Digital Services Act (DSA) in order to hold online platforms accountable; 2. technology, through the use of Artificial Intelligence aimed at monitoring and preventing online threats; and 3. education, focused on promoting digital awareness and the use of legal design to make legal rules accessible and comprehensible to younger generations. The objective is to ensure a safe and inclusive digital future through the synergy between institutions, technology, and education.

Questo contributo esprime una breve considerazione circa la protezione di minorenni e dei soggetti fragili destinatari dei rischi della giungla digitale, come cyberbullismo e catfishing. L'analisi propone un approccio olistico basato su tre pilastri: 1. quadro normativo per l'implementazione di sistemi di verifica dell'età e applicazione del Digital Services Act (DSA) per responsabilizzare le piattaforme; 2. tecnologia per l'uso dell'Intelligenza Artificiale volto a monitorare e prevenire minacce online; 3. educazione verso la promozione della con-

*[1] Irene Coppola, research fellow at the Suor Orsola Benincasa University in Naples and Senior Researcher Research Centre in European Private Law (ReCEPL) §§ 1,2,3,7,8. The entire text has been translated into english by Irene Coppola.

*[2] Cristina Mingo, lawyer and voluntary collaborator Research Centre in European Private Law University Studi Suor Orsola Benincasa, §§ 4, 5, 6.

[§] This contributiond evelops the presentationat the conference 'Cybersecurity e prodotti: sfide e responsabilità nelle società interconnesse', 29 November 2024, Italian Supreme Court.

sapevolezza digitale e l'uso del legal design per rendere le norme accessibili e comprensibili alle nuove generazioni. L'obiettivo è garantire un futuro digitale sicuro e inclusivo attraverso la sinergia tra istituzioni, tecnologia e formazione.

1. Introduction

Technological progress and the growing interconnectedness of our societies have made Cybersecurity a central issue in contemporary debate. Whilst digitalisation offers countless opportunities for development and communication, it also exposes individuals and communities to unprecedented and complex risks. Particular attention must be paid to protecting the most vulnerable, such as children and vulnerable adults, who find themselves operating in a constantly evolving digital environment, often without the tools to recognise and deal with the dangers. This article aims to explore current challenges and future scenarios in the field of Cybersecurity, with a specific focus on the protection of children and vulnerable people, by touching upon the regulatory framework, emerging technologies such as Artificial Intelligence, and the crucial importance of digital education[1].

2. Cybersecurity challenges in organisations

General overview

Modern societies are intrinsically interconnected, with an ever-increasing reliance on digital infrastructure. This interdependence creates fertile ground for new types of cyber threats. Cybersecurity challenges no longer concern only the protection of data or networks, but extend to the security of digital products and services, with direct implications for daily life and the global economy. The complexity of these scenarios requires a multidisciplinary approach and the constant evolution of defence strategies[2].

Cybersecurity and protection of minors and vulnerable people

A critical aspect of Cybersecurity concerns the protection of vulnerable groups. Minors and vulnerable individuals are often less aware of online

[1] Agenzia per la Cybersicurezza Nazionale (ACN): *Strategia Nazionale di Cybersicurezza 2022-2026.* È il documento fondamentale che definisce gli obiettivi di resilienza digitale dell'Italia. Link: https://www.acn.gov.it/

[2] P.L. Montessoro, *Cybersecurity and the protection of vulnerable individuals in the digital age: challenges and prevention strategies*, in «Istituzioni del Federalismo», (2019), p. 783 ss.

risks and are more easily exposed to phenomena such as cyberbullying, grooming, and fraud. Protecting these groups requires not only technological and regulatory interventions, but also extensive awareness-raising and training efforts aimed at promoting the informed and safe use of digital tools.

The Italian context: CLUSIT and the Global Cybersecurity Forum

In Italy, the issue of Cybersecurity is being addressed with increasing attention by various institutions and organisations. Bodies such as CLUSIT (the Italian Association for Information Security) and the Global Cybersecurity Forum actively contribute to the discussion and promotion of initiatives aimed at strengthening cyber defences at national and international level. The Department of Public Security plays a key role in ensuring online security, highlighting the importance of a coordinated approach between the public and private sectors. The international conference on 29 November 2024 at the University of San Raffaele, focusing on 'Cybersecurity and products: challenges and responsibilities in interconnected societies' and 'Cybersecurity and the protection of minors and vulnerable people', demonstrates the topicality and relevance of these issues within the academic and professional spheres[3].

3. The digital jungle: online risks

Whilst the digital world offers immense opportunities, it can also be a veritable jungle, fraught with dangers for less experienced users. Certain phenomena are particularly harmful to children and vulnerable people.

Cyberbullying

Cyberbullying is one of the most widespread and dangerous threats. It involves acts of abuse and harassment carried out via digital means, which can have devastating consequences for the psychological well-being of victims. The pervasive nature of the digital world makes it difficult to escape such attacks, which can occur 24 hours a day, 7 days a week.

Catfishing

Catfishing is another worrying phenomenon, which involves creating false online identities to deceive and manipulate victims, often for the purposes of extortion or grooming. The emotional and financial consequen-

[3] Italian Data Protection Authority (Garante Privacy): Specific focus on minors and "digital vulnerability." The website offers measures and guidance (es. *Cyberbullismo, Revenge Porn, Social Network*). Link: https://www.garanteprivacy.it/minori.

ces of catfishing can be very serious, especially when the victims are young or particularly vulnerable[4].

4. Regulatory developments in online protection

The growing prevalence of online risks has prompted legislators to develop an increasingly comprehensive regulatory framework[5].

The laws aim to establish liability, define cybercrimes and implement protective measures. However, the rapid pace at which the digital landscape evolves makes regulatory adaptation a constant challenge. A crucial aspect of protecting children online is the implementation of effective age verification systems. Such systems must meet strict requirements to ensure both protection and accessibility:

– Proportionality: the instrument must be as non-invasive as possible.

– Data protection: the system must comply with the principles of the GDPR, in particular data minimisation and privacy by design and by default.

– Involvement of independent third parties: user identification through age verification should ideally involve impartial third parties.

– Security: adequate Cybersecurity measures must be put in place.

– Accuracy and effectiveness: the system must minimise the possibility of error.

– Functionality, ease of use and no barriers to accessing online content.

– Inclusivity and non-discrimination

– Transparency: clear and comprehensive explanations of how the system works must be provided, and data on its accuracy and effectiveness must be made available.

– Training and information: it is essential to promote awareness campaigns on the risks associated with the internet.

– Complaints handling: a procedure must be put in place for handling complaints in the event of incorrect age determinations.

[4] Link:https://www.enisa.europa.eu/publications/enisa-threat-landscape-2024.

[5] Law 71/2017: provisions to protect minors in the prevention and fight against cyberbullying.

Link: https://www.normattiva.it/uri-res/N2Ls?urn:nir:stato:legge:2017-05-29;71EU AI Act (Regolamento IA): new European rules place specific restrictions on AI systems that can manipulate vulnerable individuals.

Link:https://www.europarl.europa.eu/topics/it/article/20230601ST093804/normativa-sull-ia-la-prima-regolamentazione-sull-intelligenza-artificiale

Convenzione di Budapest: on cybercrime (Council of Europe). Link: https://www.coe.int/it/web/portal/coe-action-against-cybercrime.

– These general requirements that an age verification system must meet highlight the complexity and sensitivity of the issue, striking a balance between the need for protection and users' fundamental rights[6].

5. The role of Digital Services Act (DSA)

The European Union's Digital Services Act (DSA)[7] is a cornerstone of digital services regulation. This legislation introduces new responsibilities for online platforms, in particular for with regard to content moderation and user protection. Its provisions aim to create a safer and more transparent online environment, with a specific focus on safeguarding fundamental rights and combating disinformation and illegal content. The DSA is therefore an essential tool for tackling the risks outlined above and for strengthening Cybersecurity at European level.

6. Artificial Intelligence and the protection of children online

Artificial Intelligence (AI) is emerging as a promising tool in the fight against online threats, particularly when it comes to protecting minors.

The potential of AI in online security

AI can be used to identify and mitigate risks such as cyberbullying and grooming through data analysis and the detection of anomalous patterns. AI-based systems can monitor online interactions, flag inappropriate content and even predict potentially dangerous situations, providing a level of protection that traditional methods struggle to achieve.

Synergies between Artificial Intelligence and the Digital Services Act

The combination of Artificial Intelligence and the Digital Services Act (DSA) can create a particularly robust ecosystem for online protection. The DSA provides the regulatory framework that requires platforms to adopt security measures, whilst AI offers the technological tools to implement these

[6] G. Pesci, *Perspectives and recent developments in child protection on line*, in «Ciberspazio e Diritto. Rivista internazionale di informatica giuridica», II (2022), pp. 311 ss.

[7] Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a single market for digital services and amending Directive 2000/31/EC (Digital Services Regulation). Link: https://www.google.com/search?q=https://eur-lex.europa.eu/legal-content/IT/TXT/%3Furi%3DCELEX:32022R2065.

measures efficiently and effectively. This synergy is crucial for tackling the complexity and scale of cyber threats in the digital age[8].

7. The importance of Cybersecurity education: the case of Iris Stein

In addition to technological and regulatory solutions, education plays an irreplaceable role in promoting the safe and responsible use of digital technology. The "Iris Stein case" is a prime example of the importance of educating young people about Cybersecurity[9]. These real-life examples serve to illustrate the actual consequences of online dangers and to highlight the need to equip young people with the critical skills required to navigate cyberspace safely. Awareness of the risks, combined with knowledge of best security practices, is essential for fostering responsible digital citizenship[10].

8. Conclusions and future prospects: legal design for the teaching of private law

Protecting vulnerable individuals in the digital environment requires a holistic approach that combines effective legislation, advanced technologies and, above all, sound education. An innovative approach aims to make the complexities of the law, including Cybersecurity legislation, more accessible and understandable, particularly for younger generations. Through clear and intuitive educational tools, legal design can help foster more aware and protected digital citizens through alert mechanisms featuring intuitive and clear images and infographics. In an age where offline and online life are increasingly interconnected, the ability to protect children and vulnerable people from the risks of the digital world becomes an ethical and social imperative. The challenges are many, but the joint commitment of institutions, academics, technologists and

[8] A. Gorgoni, *Vulnerability in the digital environment and the protection of freedom of will,* in«Rivista Persona E Mercato», III (2024), pp. 915 ss.

[9] The Iris Stein case is a fictional investigative simulation used in educational settings (such as school hackathons) to teach Cybersecurity and data protection to young people .

Link: https://www.agendadigitale.eu/sicurezza/il-caso-di-iris-stein-limportanza-di-educare-gli-studenti-alla-cybersecurity/.

[10] D. Alfieri, *Internet: when the "net" captures minors Internet,* in «Rivista italiana di Informatica e Diritto»,I (2022), pp. 53 ss.; according to the Author, digital education must be addressed in a comprehensive manner, from legislators to educators, the latter often too slow to keep up with the frenzy of the internet to which the new generations are subjected.

families, combined with targeted education and a robust regulatory framework, can build a safer and more inclusive digital future[11].

[11] C. D'Onofrio, *Legal Design and Artificial Intelligence: New prospects for the future of law*, in «i-lex – Rivista di Scienze Giuridiche, Scienze Cognitive e Intelligenza Artificiale», I(2023), pp. 26 ss.; Legal Design and artificial intelligence are tools at the service of law that have the intrinsic capacity to maximize its comprehensibility, usability – and therefore accessibility – and efficiency; A. Fabiano, Artificial Intelligence and Teaching Processes: A New Educational Paradigm, (2025), pp. 34 ss; a vision of education through gamification appears to be a new push towards the use of technology, especially for the benefit of minors; in this regard F. Abbes, S. Bennani, A.Maalel, *Generative AI and Gamification for Personalized Learning: Literature Review and Future Challenges*, in «SN Computer Science», VIII (2024), pp. 1154 ss.

BIBLIOGRAPHY

Abbes F., Bennani S., Maalelm A., *Generative AI and Gamification for Personalized Learning: Literature Review and Future Challenges*, in «SN Computer Science», VIII (2024), pp.1154 ss.

Alfieri D., *The Internet: when the 'web' ensnares minors*, in «Rivista italiana di Informatica e Diritto», I (2022), p. 53 ss.

D'Onofrio C., *Legal Design e Intelligenza Artificiale: Nuove prospettive per il futuro del diritto,* in *«i*-lex – Rivista di Scienze Giuridiche, Scienze Cognitive e Intelligenza Artificiale», I (2023), pp. 26–36.

Fabiano A., *Artificial Intelligence and teaching processes. A new educational paradigm*, (2025), pp. 34 ss.

Gorgoni A., *La vulnerabilità nell'ambiente digitale e la protezione della libertà del volere*, in «Persona e mercato», III (2024), pp. 915 ss.

Montessoro P.L., *La cybersecurity and the protection of vulnerable individuals in the digital age: challenges and prevention strategies,* in «Istituzioni del Federalismo», (2019), pp. 783 ss.

Pesci G., *Outlook and recent developments in the protection of children online, in* «Ciberspazio e Diritto, Rivista internazionale di informatica giuridica», II (2022) pp. 311 ss.

Pileggi B., *Children online: a constitutional issue*, in «BioLaw Journal-Rivista di BioDiritto», I (2024), pp. 361 ss.

GIUSEPPE CORASANITI*

Cybersecurity and medical devices: a progressive approach in Digital bioethics[§]

Abstract

This contribution examines how the digital transformation of healthcare, from the Internet of Medical Things and AI assisted diagnostics to the pervasive use of health big data, requires the maturation of digital bioethics as a normative and operational framework capable of aligning technological innovation with human dignity, patient safety, and data protection. It grounds this claim in the supranational architecture of fundamental rights (EU Charter; ECHR) and in the Oviedo Convention, whose cardinal principles — primacy of the person, informed consent, and protection of private life — are technologically refracted in the GDPR and the AI Act. Within the EU legal order, the paper maps a multilayered regime in which the MDR, Cybersecurity Act, and NIS2 interact with data protection to structure safety, performance, and cyber resilience obligations for connected medical devices. Jurisprudence of the CJEU — on proportionality, necessity, surveillance limits, liability, and human oversight in automated processes — operates as an interpretive compass for Cybersecurity by design and risk based compliance. In parallel, the ECHR Article 2 right to life articulates States' positive and procedural obligations to regulate, monitor, and investigate device related risks, especially for life sustaining technologies. The article concludes by proposing codes of conduct as a practical vehicle to translate digital bioethics into enforceable governance, ensuring transparency, accountability, non discrimination, and

* Full Professor of Methodology and Logic of Digital Law, Legal Informatics, and Digital Ethics at Mercatorum University in Rome.

§ This contribution develops the presentation at the conference 'Cybersecurity e prodotti: sfide e responsabilità nelle società interconnesse', 29 November 2024, Italian Supreme Court.

sustainability, thereby reconciling innovation with fundamental rights in an increasingly connected clinical ecosystem.

Il presente contributo esamina come la trasformazione digitale dell'assistenza sanitaria, dall'Internet of Medical Things e dalla diagnostica assistita dall'Intelligenza Artificiale, fino all'uso pervasivo dei big data sanitari, richieda la maturazione di una bioetica digitale quale cornice normativa e operativa capace di allineare l'innovazione tecnologica con la dignità umana, la sicurezza del paziente e la protezione dei dati. Tale tesi è fondata nell'architettura sovranazionale dei diritti fondamentali (Carta dei diritti dell'UE; CEDU) e nella Convenzione di Oviedo, i cui principi cardine — primato della persona, consenso informato e tutela della vita privata — trovano oggi una declinazione tecnologica nel GDPR e nell'AI Act. Nell'ordinamento dell'Unione, il lavoro mappa un regime multilivello in cui il MDR, il Cybersecurity Act e la Direttiva NIS2 interagiscono con la protezione dei dati per strutturare obblighi di sicurezza, prestazione e cyber resilience dei dispositivi medici connessi. La giurisprudenza della Corte di giustizia dell'Unione europea — sui principi di proporzionalità e necessità, sui limiti alla sorveglianza, sulla responsabilità del produttore e sulla supervisione umana nei processi automatizzati — funge da bussola interpretativa per il Cybersecurity by design e per la compliance basata sul rischio. In parallelo, l'art. 2 CEDU sul diritto alla vita articola obblighi positivi e procedurali in capo agli Stati in ordine alla regolazione, al monitoraggio e all'investigazione dei rischi connessi ai dispositivi, in particolare quelli salva vita. L'articolo conclude proponendo i codici di condotta come veicolo pratico per tradurre la bioetica digitale in una governance effettiva, garantendo trasparenza, responsabilità, non discriminazione e sostenibilità, così da conciliare innovazione e diritti fondamentali in un ecosistema clinico sempre più connesso.

1. Introduction and Foundations of Digital Bioethics

Digital innovation applied to medicine — from connected medical devices (Internet of Medical Things) to Artificial Intelligence for diagnosis, up to the massive use of health big data — has radically transformed clinical, organizational and research dynamics. This transformation, while bringing benefits in terms of effectiveness, precision and personalization of care, nevertheless introduces ethical and legal risks that go far beyond traditional medical bioethics. Hence the need to develop digital bioethics: a corpus of principles, norms and operational practices aimed at ensuring that technological innovation

develops in harmony with the fundamental rights of the person, patient safety and the protection of personal data[1].

The legal foundations of digital bioethics are found in international and supranational law and in the Charter of Fundamental Rights of the European Union. In particular, art. 1 (human dignity), art. 3 (right to the integrity of the person) and art. 8 (protection of personal data) . Thus art. 8 of the European Convention on Human Rights (ECHR) establishes the right to respect for private and family life. But these statements should perhaps be considered in the light of the Oviedo Convention (1997), the protection of human rights and the dignity of the human being with regard to the applications of biology and medicine. The *Convention for the Protection of Human Rights and the Dignity of the Human Being with regard to the Applications of Biology and Medicine* — known as the Oviedo Convention, signed on 4 April 1997 under the auspices of the Council of Europe — is the first legally binding international treaty on bioethics and biomedicine. Italy ratified it with Law No. 145 of 28 March 2001. Within the framework of the Convention (art. 1) the principle of the protection of every human being in his dignity and identity is established and guaranteeing everyone, without discrimination, respect for his integrity and other fundamental rights and freedoms with regard to the applications of biology and medicine. Thus, in art. 2 the "primacy" of the human being is posited by defining that "*the interest and good of the human being must prevail over the sole interest of society or science." Thus art. 5 which regulates the principle of informed consent to health treatment with the right to "adequate information about the purpose and nature of the intervention, as well as its consequences and risks. In the same context, rt. 10 (Privacy and right to information) which provides that "Everyone has the right to respect for his or her private life when it comes to information relating to his or her health. "Everyone has the right to know any information collected about their health. A person's will not to be informed must be respected." Art. 16 on the Protection of persons who lend themselves to research as no research can be conducted on a person without all the following conditions being met: a) there is no alternative to research on human beings; (b) the risks to which the person is exposed are not disproportionate to the potential benefits; (c) the research is carried out under the supervision of a competent authority[2]".*

[1] T. Magnin, P. Giorgini, *Digital Driven Technosciences: Epistemological and Ethical Questioning,* in «Etica & politica-Ethics & politics», XXIV, I (2022), p. 379.

[2] F. Cirillo, *The Impact of e-Health on Privacy and Fundamental Rights: From Confidentiality to Data Protection Regulation*, in «European Journal of Privacy Law & Technologies», II (2019), p. 95.

The Oviedo Convention thus constitutes the international ethical-legal basis on which reflection on digital bioethics is grafted today, because it sets out some cardinal principles such as the primacy of the person over the interest of science (art. 2), the centrality of free and informed consent (art. 5) and finally the protection of private life and health information (art. 10). These same principles are found today, declined in a technological key, in the GDPR (Articles 9, 25 and 32) and in the AI Act, to guarantee that the use of digital medical devices and Artificial Intelligence does not sacrifice the dignity and freedom of the person.

2. EU Regulatory Framework: MDR, GDPR, NIS2, Cybersecurity Act

The intersection between medical devices and Cybersecurity represents one of the most complex and rapidly evolving areas of EU law, requiring careful analysis of multiple regulatory frameworks and fundamental principles of Union law. This comprehensive examination addresses the legal challenges posed by the digitalization of healthcare technologies within the European legal order. The European Union has developed a sophisticated multi-layered approach to regulating medical devices and Cybersecurity, reflecting the cross-cutting nature of these issues across various policy domains. The primary legal foundation rests on Article 114 TFEU, which provides the competence for harmonization measures necessary for the establishment and functioning of the internal market, and Article 168 TFEU, which addresses public health protection.

The Medical Device Regulation (MDR) 2017/745 represents the cornerstone of the current regulatory framework, replacing the previous directive-based system with a more comprehensive and uniform approach. This regulation, fully applicable since May 2021, establishes enhanced requirements for medical device safety and performance, including specific provisions addressing software and connectivity aspects. The regulation operates on the principle of total harmonization, leaving minimal room for national variations in essential requirements.

Parallel to the MDR, the Cybersecurity Act (Regulation 2019/881) establishes the European Union Agency for Cybersecurity (ENISA) and creates a framework for European Cybersecurity certification schemes. While not specifically targeting medical devices, this regulation provides horizontal Cybersecurity principles applicable across sectors. The Network and Information Security Directive (NIS Directive 2016/1148), recently updated by NIS2

Directive 2022/2555, extends Cybersecurity obligations to essential services, potentially including healthcare providers using connected medical devices. The General Data Protection Regulation (GDPR) adds another crucial layer, particularly relevant for connected medical devices processing personal health data. The intersection between medical device regulation and data protection creates complex compliance obligations, especially regarding consent mechanisms, data minimization principles, and cross-border data transfers in clinical settings.

3. CJEU Jurisprudence: Proportionality, Liability, and AI Oversight

The Court of Justice of the European Union has developed significant jurisprudence relevant to medical devices and Cybersecurity, though often in fragmented contexts requiring synthesis and analogical reasoning. The landmark case *Digital Rights Ireland* (C-293/12)[3] established fundamental principles regarding the balance between technological capabilities and fundamental rights, particularly relevant for connected medical devices collecting vast amounts of personal health data. In *Tele2 Sverige* (C-203/15)[4], the Court clarified the scope of surveillance obligations and the necessity principle in telecommunications, providing guidance applicable to medical device connectivity requirements. The Court's emphasis on proportionality and necessity in technological surveillance measures translates directly to Cybersecurity requirements for medical devices, particularly regarding data collection and monitoring capabilities.

The *Boston Scientific case* (C-503/13)[5] addressed medical device liability and the interpretation of harmonized standards, establishing important precedents for the relationship between technical specifications and legal obligations. The Court emphasized that compliance with harmonized standards creates a presumption of conformity but does not exclude manufacturer liability for defects, a principle particularly relevant for Cybersecurity vulnerabilities in medical devices.

More recently, the Court's jurisprudence on Artificial Intelligence and automated decision-making, while not specifically addressing medical devices, provides valuable interpretative guidance for AI-enabled medical technologies. The Court has consistently emphasized the importance of human oversight and the need for transparent, explainable automated processes, principles directly applicable to AI-driven medical diagnostics and treatment

[3] https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex:62012CJ0293

[4] https://curia.europa.eu/juris/document/document.jsf?docid=186492&doclang=IT

[5] https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:62013CJ0503

devices. The principle of supremacy of EU law operates with particular complexity in the medical devices Cybersecurity context, as it intersects with national healthcare systems and security policies traditionally considered within State competence. The Court has established that while Member States retain competence over healthcare organization, they must exercise this competence in compliance with EU law, including Cybersecurity and medical device regulations. Direct effect manifests differently across the various legal instruments applicable to medical devices and Cybersecurity. While the MDR, being a regulation, enjoys direct applicability in all Member States, its specific Cybersecurity provisions often require implementation through technical standards and guidance documents. The GDPR similarly applies directly but requires coordination with national data protection authorities for enforcement.

The *proportionality* principle assumes crucial importance in balancing Cybersecurity requirements with medical device functionality and patient care continuity. The Court has established that Cybersecurity measures must be appropriate to the risk level and not unduly burden healthcare providers or impede patient access to medical care. This principle guides the interpretation of "state of the art" Cybersecurity requirements under the MDR.

Subsidiarity plays in some time a complex role, as Cybersecurity threats are inherently cross-border while healthcare delivery remains primarily national. The Court has recognized that effective Cybersecurity for medical devices requires EU-level coordination while respecting national healthcare system organization. This creates a nuanced application of subsidiarity that favors EU action for technical standards and threat intelligence while preserving national competence for healthcare delivery organization. The relationship between EU medical device Cybersecurity law and national legal systems creates multiple layers of complexity. Member States must transpose and implement EU requirements while maintaining their constitutional obligations regarding healthcare provision and national security. The Court has established that national authorities cannot invoke national security exceptions to circumvent EU medical device safety requirements, but they retain competence to address specific national vulnerabilities.

National courts play a crucial role in applying EU principles to specific cases involving medical device Cybersecurity incidents. The preliminary reference procedure under Article 267 TFEU provides the mechanism for ensuring uniform interpretation, though the technical complexity of Cybersecurity issues sometimes challenges traditional judicial competences. The Court has encouraged national courts to seek expert technical advice while maintaining judicial independence in legal interpretation.

The enforcement mechanisms vary significantly across Member States, reflecting different administrative traditions and technical capabilities. Some Member States have established specialized Cybersecurity authorities with medical device competence, while others rely on traditional medical device notified bodies to assess Cybersecurity compliance. This variation creates potential for regulatory arbitrage and uneven protection levels across the internal market.

Member State notification obligations under various EU instruments create complex reporting requirements for Cybersecurity incidents affecting medical devices. The interaction between NIS Directive notification requirements, GDPR breach notification obligations, and MDR incident reporting creates potential for conflicting or overlapping obligations requiring careful coordination.

The rapid pace of technological development in medical devices creates ongoing challenges for the regulatory framework. Connected medical devices, Artificial Intelligence applications, and Internet of Medical Things (IoMT) ecosystems evolve faster than traditional regulatory processes can adapt. This creates tension between innovation promotion and risk management that requires flexible legal interpretation.

Cross-border healthcare provision using connected medical devices raises complex questions about applicable law and enforcement jurisdiction. A medical device manufactured in one Member State, certified by a notified body in another, and used to treat patients in a third, potentially storing data in a fourth, creates unprecedented jurisdictional complexity requiring careful application of conflict of laws principles.

The international dimension adds further complexity, as medical device Cybersecurity often involves third-country manufacturers, cloud services, and threat actors. The Court's jurisprudence on adequacy decisions under GDPR provides some guidance, but the specific requirements for medical device data processing require further development.

The EU institutions have begun developing more integrated approaches to medical device Cybersecurity. The European Commission's proposed Cyber Resilience Act aims to establish horizontal Cybersecurity requirements for connected products, potentially including medical devices. This creates opportunities for better coordination between sectoral and horizontal Cybersecurity requirements. The development of common technical standards through European standardization organizations provides a mechanism for translating legal requirements into practical technical specifications. The Court has recognized the importance of these standards while maintaining that they cannot replace legal obligations or limit fundamental rights protection.

Enhanced cooperation between regulatory authorities offers pathways for better coordination. The Medical Device Coordination Group established under the MDR provides a forum for addressing Cybersecurity issues, while ENISA's sectoral expertise can inform healthcare-specific guidance. The Court's jurisprudence on administrative cooperation supports these collaborative approaches.

The establishment of EU-wide Cybersecurity certification schemes for medical devices could provide greater legal certainty and market harmonization. Such schemes would need to respect the principle of mutual recognition while ensuring adequate protection levels across all Member States.

Healthcare providers must navigate complex compliance obligations spanning multiple EU legal instruments. The Court's emphasis on practical effectiveness suggests that compliance frameworks should integrate Cybersecurity requirements into existing quality management systems rather than creating parallel obligations. This integrated approach aligns with the risk-based methodology emphasized in CJEU jurisprudence. Manufacturer obligations extend beyond initial market placement to ongoing Cybersecurity monitoring and update provision. The Court's interpretation of producer liability principles suggests that manufacturers cannot disclaim responsibility for Cybersecurity vulnerabilities discovered after market placement, particularly when updates are technically feasible.

Cross-border incident response requires coordination between multiple national authorities and EU agencies. The Court's jurisprudence on administrative cooperation supports streamlined procedures while maintaining national competences for emergency response. Practical protocols should anticipate the need for rapid information sharing while respecting patient confidentiality and commercial sensitivity.

The evolving nature of Cybersecurity threats requires adaptive legal interpretation that accommodating technological developments. The Court's living instrument approach to EU law interpretation supports evolutionary interpretation of Cybersecurity requirements in response to emerging threats and technological capabilities.

Artificial Intelligence integration in medical devices will require careful application of fundamental rights principles, particularly regarding transparency, explainability, and human oversight. The Court's emerging jurisprudence on automated decision-making provides interpretative guidance that will become increasingly relevant as AI-enabled medical devices proliferate.

The development of digital health spaces and European Health Data Space initiatives will create new contexts for applying medical device Cyberse-

curity law. These developments require coordinated interpretation of multiple legal instruments while maintaining the coherence of the overall EU legal order.

In conclusion, the regulation of medical devices Cybersecurity within EU law represents a complex intersection of multiple regulatory frameworks, fundamental principles, and evolving technological realities. The Court of Justice's interpretative approach emphasizes practical effectiveness, fundamental rights protection, and internal market integration while respecting Member State competences in healthcare organization. Future developments will require continued judicial interpretation that balances innovation promotion with risk management, maintaining the coherence and effectiveness of the European legal order in this critical area of public health and technological development. The success of this regulatory approach depends on continued cooperation between EU institutions, national authorities, healthcare providers, and industry stakeholders, guided by the fundamental principles established in CJEU jurisprudence and adapted to the evolving challenges of medical device Cybersecurity in an increasingly connected healthcare environment.

4. ECHR Article 2: Right to Life and State Obligations for Medical Technologies

The intersection between medical devices and the right to life under Article 2 of the European Convention on Human Rights presents one of the most compelling contemporary challenges in European human rights law. As healthcare increasingly relies on sophisticated medical technologies, the Strasbourg Court faces complex questions about State obligations to regulate, monitor, and ensure the safety of life-sustaining medical devices within the framework of fundamental rights protection.

Article 2 of the Convention, enshrining the right to life, operates as the foundation of the entire human rights system, described by the European Court as the most fundamental provision from which no derogation is permitted even in times of war or public emergency. In the context of medical devices, this absolute character creates particular interpretative challenges when balancing competing interests such as innovation, accessibility, and safety. The primary obligation under Article 2 requires States to refrain from intentionally taking life, except in strictly defined circumstances. However, the Court's jurisprudence has evolved to recognize extensive positive obligations requiring States to take appropriate steps to safeguard life. In the medical device

context, this evolution proves crucial as it extends State responsibility beyond direct action to encompass regulatory oversight and protective measures.

Article 8, protecting private and family life, frequently intersects with Article 2 in medical device cases, particularly regarding decisions about treatment withdrawal or the use of life-sustaining technologies. The Court has established that individuals retain autonomy over medical decisions affecting their own lives, creating tension between State obligations to protect life and respect for personal autonomy.

Protocol 12's general prohibition of discrimination potentially amplifies these issues when medical device access varies across populations or when technological solutions are not equally available to all patients requiring life-sustaining treatment.

However, the limited ratification of Protocol 12 means that discrimination claims typically must be anchored to substantive Convention rights. The European Court's approach to medical cases has developed through several landmark decisions that establish the framework for analyzing medical device issues. The *Pretty v. United Kingdom* case (2002)[6] established fundamental principles regarding State obligations in medical contexts, clarifying that Article 2 does not create a right to die but recognizes individual autonomy in medical decision-making under Article 8. In *Lambert and Others v. France* (2015)[7] the Grand Chamber addressed life-sustaining medical treatment, establishing that decisions to withdraw or continue treatment must follow careful procedural safeguards while respecting patient autonomy or, when impossible, best interests determinations.

This jurisprudence directly applies to medical device cases where devices maintain life functions. The Court's decision in *Oyal v. Turkey* (2010)[8]addressed State obligations regarding medical care quality, establishing that States must ensure adequate healthcare systems capable of protecting life. While not specifically addressing medical devices, this precedent creates obligations for regulatory oversight of medical technologies essential to life preservation.

More recently, the Court has begun addressing technological aspects of healthcare in cases like M.A. v. Denmark, (2018)[9] where medical record systems and healthcare delivery methods intersected with fundamental rights. These cases suggest the Court's readiness to apply traditional human rights principles to technological healthcare contexts.

[6] https://hudoc.echr.coe.int/fre#{%22itemid%22:[%22001-60448%22]}
[7] https://hudoc.echr.coe.int/eng#{%22itemid%22:[%22002-10758%22]}
[8] https://hudoc.echr.coe.int/eng#{%22itemid%22:[%22001-92945%22]}
[9] https://hudoc.echr.coe.int/eng#{%22itemid%22:[%22001-211178%22]}

The Calvelli and Ciglio v. Italy (2002)[10] judgment established important principles regarding State obligations in medical negligence cases, clarifying that gross negligence in medical care can engage Article 2. This principle extends to medical device failures when such failures result from inadequate regulatory oversight or systematic deficiencies in safety monitoring.

The European Court has established that Article 2 requires States to establish effective regulatory frameworks for protecting life, including in medical contexts. For medical devices, this translates into positive obligations to create and maintain adequate regulatory systems ensuring device safety, efficacy, and appropriate use in life-threatening situations. The Court recognizes different levels of State obligation depending on the circumstances. When medical devices are used in State-controlled healthcare settings, the obligation is more stringent than when devices are used in private healthcare contexts. However, even private healthcare requires adequate regulatory oversight to ensure life protection standards. .The margin of appreciation in medical device regulation acknowledges that States have primary responsibility for organizing healthcare systems while maintaining minimum protection standards. This margin is narrower when fundamental life interests are at stake, requiring States to demonstrate that their regulatory choices adequately protect the right to life.

Procedural obligations under Article 2 extend to medical device contexts, requiring effective investigation of incidents where device failures contribute to loss of life. The Court has established that investigations must be capable of leading to identification and punishment of those responsible, including regulatory authorities if systematic failures enabled preventable deaths.

The prevention principle requires States to take reasonable measures to prevent loss of life when authorities knew or ought to have known of real and immediate risk. In medical device contexts, this could include obligations to recall dangerous devices, issue safety warnings, or implement enhanced monitoring for high-risk technologies.

Negative obligations under Article 2 require States to refrain from actions that would deprive individuals of life. In medical device contexts, this extends to regulatory decisions that knowingly permit dangerous devices to remain in use or that inadequately assess life-threatening risks. The prohibition is not limited to direct action but encompasses regulatory negligence that foreseeably results in loss of life.

The Court has clarified that negative obligations extend to situations where State agents or State-controlled institutions use medical devices. Hospitals, emergency services, and other public healthcare providers must ensure

[10] https://hudoc.echr.coe.int/eng#{%22itemid%22:[%22001-60329%22]}

that life-sustaining medical devices meet appropriate safety standards and are properly maintained and operated.

Positive obligations require States to establish effective regulatory frameworks for medical device safety. This includes pre-market authorization systems, post-market surveillance mechanisms, and adequate enforcement procedures. The Court has recognized that modern healthcare's complexity requires sophisticated regulatory responses proportionate to the life risks involved.

The institutional framework obligation requires States to establish competent authorities with adequate resources and expertise to regulate medical devices effectively. The Court's jurisprudence suggests that understaffed or technically inadequate regulatory bodies could violate Article 2 when their deficiencies contribute to preventable deaths.

Information obligations require States to ensure that healthcare providers and patients receive adequate information about medical device risks and benefits. The Court has recognized that informed decision-making in medical contexts is essential to respecting both life protection and personal autonomy.

International cooperation obligations have emerged in the Court's recent jurisprudence, recognizing that medical device safety often requires cross-border coordination. States cannot use territorial limitations to excuse failures in medical device oversight when international cooperation could prevent loss of life.

Article 2's procedural dimension requires effective official investigation when loss of life occurs in circumstances potentially engaging State responsibility. For medical device-related deaths, this creates obligations for thorough investigation of device failures, regulatory oversight adequacy, and institutional responses to safety concerns. The investigation must be capable of leading to identification and punishment of those responsible, whether individual healthcare providers, device manufacturers, or regulatory authorities. The Court has emphasized that investigations cannot be merely formal but must genuinely seek to determine whether deaths were preventable through better regulatory oversight.

Independence requirements mean that investigations of medical device-related deaths must be conducted by authorities independent of those potentially responsible for regulatory failures. This may require specialized investigative bodies or external oversight when regulatory authorities' decisions are questioned.

Public scrutiny elements require that families and the public have adequate access to investigation findings, subject to legitimate confidentiality interests. The Court has recognized that transparency in medical device

safety investigations serves both individual justice and public safety interests. Reasonable expedition in investigations acknowledges the urgency of addressing ongoing safety risks while ensuring thoroughness. The Court balances the need for quick action to prevent further deaths against the requirement for comprehensive investigation of complex technical issues.

The Court frequently faces situations where the right to life intersects with other Convention rights in medical device contexts. Patient autonomy under Article 8 may conflict with State obligations to preserve life, particularly regarding decisions to discontinue life-sustaining device treatment. The Court has also established that competent adults have the right to refuse life-sustaining treatment, including medical devices, even when such refusal will result in death. However, States retain obligations to ensure that such decisions are genuinely autonomous and made with adequate information about consequences and alternatives.

For incompetent patients, the Court requires procedural safeguards ensuring that decisions about life-sustaining medical devices are made in the patient's best interests following appropriate consultation with family members and medical professionals. The specific procedures may vary among States within their margin of appreciation.

Religious and cultural considerations may influence how individuals and families approach medical device use, particularly for life-sustaining technologies. The Court has recognized that Article 9 protections may be relevant while maintaining that religious beliefs cannot override fundamental life protection obligations when individuals lack capacity for autonomous decision-making. Economic considerations cannot justify failing to provide life-sustaining medical devices when needed, though the Court recognizes that resource limitations may influence how States organize their healthcare systems. The key principle is that decisions must be based on medical need rather than discriminatory factors.

Individual remedies in medical device cases under Article 2 typically focus on just satisfaction for families of deceased victims, as restitution is impossible. The Court has awarded both pecuniary damages for financial losses and non-pecuniary damages for suffering caused by the loss of life and procedural violations. The Court increasingly recognizes that monetary compensation alone may be insufficient when systematic regulatory failures contribute to medical device-related deaths. Declaratory relief acknowledging State responsibility can provide important vindication for families while highlighting systemic problems requiring broader solutions.

General measures required following Article 2 violations often involve substantial changes to regulatory systems. The Court may require States

to strengthen medical device oversight, improve post-market surveillance, or enhance healthcare provider training on device safety.

Legislative reforms may be necessary when existing laws inadequately protect the right to life in medical device contexts. The Committee of Ministers supervises implementation of such reforms, ensuring that changes adequately address the violations identified by the Court.

Training and capacity building requirements may emerge when violations result from inadequate expertise within regulatory authorities or healthcare institutions. The Court recognizes that effective medical device oversight requires specialized technical knowledge that States must ensure is available.

Artificial Intelligence in medical devices presents novel challenges for applying Article 2, particularly regarding accountability when AI systems make life-affecting decisions. The Court will likely need to develop jurisprudence addressing human oversight requirements and liability allocation in AI-assisted medical interventions[11].

Cybersecurity vulnerabilities in connected medical devices create new categories of life-threatening risks requiring State attention. The Court may need to address whether States have obligations to ensure Cybersecurity standards for life-sustaining medical technologies. Cross-border healthcare using medical devices raises jurisdictional questions about which State bears primary responsibility for ensuring Article 2 compliance. The Court's developing jurisprudence on extraterritorial application may prove relevant as medical device technology enables treatment across national boundaries. Personalized medicine and genetic technologies may create new forms of inequality in access to life-sustaining treatments, potentially engaging discrimination principles alongside Article 2. The Court will need to address whether States must ensure equal access to emerging medical technologies essential for life preservation.

Healthcare institutions must establish clear protocols for medical device use that respect both life protection obligations and patient autonomy. These protocols should address emergency situations, informed consent procedures, and decision-making for incompetent patients.

Regulatory authorities should develop comprehensive post-market surveillance systems capable of detecting life-threatening device failures quickly enough to prevent additional deaths. The Court's emphasis on prevention suggests that reactive approaches alone may be insufficient.

Training programs for healthcare providers must address both technical aspects of medical device operation and human rights principles governing

[11] G. Capilli, *Innovazione tecnologica e responsabilità: un breve sguardo alla "sanità mobile" (mHealth) I Technological innovation and accountability: a brief look at 'mobile health' (mHealth)*, in «European Journal of Privacy Law & Technologies», I (2023), p. 55.

their use. The intersection of technology and rights requires interdisciplinary understanding among medical professionals[12].

International cooperation mechanisms should be strengthened to address the global nature of medical device markets while maintaining accountability under national human rights obligations. States cannot use international complexity as an excuse for inadequate life protection.

In conclusion, the application of Article 2 ECHR to medical devices reflects the dynamic evolution of human rights law in response to technological advancement. The Strasbourg Court's approach emphasizes that fundamental rights principles must adapt to new contexts while maintaining their essential protective character. As medical technology continues to evolve, the European human rights system provides a robust framework for ensuring that innovation serves human dignity and life protection rather than undermining these fundamental values.

The success of this approach depends on continued dialogue between human rights advocates, medical professionals, regulatory authorities, and technology developers, guided by the Court's evolving jurisprudence and committed to the practical effectiveness that characterizes the European human rights system at its best.

5. Digital Bioethics Applied: Codes of Conduct and Ethical Governance

A digital bioethics that can be translated into codes of conduct for medical devices is fundamentally important for ensuring an ethical approach to technological innovation in healthcare. In particular:

Digital bioethics addresses new ethical challenges[13] arising from the use of digital technologies, such as Artificial Intelligence and mobile health applications, focusing on privacy, data security, patient autonomy, and transparency. This form of bioethics expands traditional principles by including new dimensions such as non-discrimination, accountability, and sustainability. In medical devices, safety, effectiveness, and respect for user rights are key

[12] F. Casarosa, *La "cybersecurity" nel quadro dello European Health Data Space,* in «Notizie di Politeia», (2025), p. 45; E. Scalcon, *Cibersicurezza e dispositivi medici: la tutela della salute e della sicurezza dei pazienti dalle vulnerabilità informatiche nel Regolamento (UE) 2017/745 e nell'Artificial Intelligence Act,* in «BioLaw Journal-Rivista di BioDiritto», CLVIII (2024), p. 183.

[13] A. Virdis, *Is it possible to define the nature of medicine? Teleological and phenomenological perspectives on medicine, its ends and its morality*, in «Medicina e morale», (2008), p. 773; G. Herranz Rodriguez, *La deontologia medica desde la tradicion hipocratica al Cristianismo,* in «Medicina e morale», (2010), p. 363, E. Villanueva Cañadas, *Derecho versus bioetica en las tomas de decisiones médicas en torno a casos limites,* in «Medicina e morale », (2011), p. 405.

ethical priorities. Codes of conduct help translate these principles into operational rules, promoting health protection, personal data security, and transparency in the technologies used. These codes are essential to deal with delicate issues such as informed consent, confidentiality, and the proper functioning of devices. Including digital bioethics in codes of conduct enables the creation of clear and shared regulatory frameworks that guide both manufacturers and healthcare professionals in the ethical use of technologies. This ensures that digital medical devices are designed, implemented, and used with respect for human dignity and legal standards, fostering human-centered and responsible care relationships.

In summary, digital bioethics translatable into codes of conduct allows the combination of technological progress and ethical protection in medical devices, ensuring safety, transparency, and respect for patients' rights in an increasingly digital context[14].

[14] M. Munzarova, *Dignity, personhood and culture: reflections on bioethics of today,* in «Medicina e morale», IV (2010), p. 547; L. Palazzani, *Dalla bioetica al biodiritto, tra teoria e prassi. Introduzione,* in «Rivista di filosofia del diritto», (2013), p. 7.

BIBLIOGRAPHY

Capilli G., *Innovazione tecnologica e responsabilità: un breve sguardo alla "sanità mobile" (mHealth) | Technological innovation and accountability: a brief look at 'mobile health' (mHealth),* in «European Journal of Privacy Law & Technologies», I (2023), p. 55.

Casarosa F., *La "cybersecurity" nel quadro dello European Health Data Space*, in «Notizie di Politeia», (2025), p. 45.

Cirillo F., *The Impact of e-Health on Privacy and Fundamental Rights: From Confidentiality to Data Protection Regulation,* in «European Journal of Privacy Law & Technologies», II (2019), p. 95.

Magnin T., Giorgini P., *Digital Driven Technosciences: Epistemological and Ethical Questioning*, in «Etica & politica-Ethics & politics», XXIV, I (2022), p. 379.

Scalcon E., *Cibersicurezza e dispositivi medici: la tutela della salute e della sicurezza dei pazienti dalle vulnerabilità informatiche nel Regolamento (UE) 2017/745 e nell'Artificial Intelligence Act,* in «BioLaw Journal-Rivista di BioDiritto», CLVIII (2024), p. 183.

STEFANO CORSO*

Technological innovation, vulnerability and ageing. The protection of the individual§

Abstract

In the multiplicity of challenges posed by technological innovation, the centrality of the concept of security emerges at the legal level and acquires new relevance through the declinations of the Cybersecurity. In this context, the language of law embraces different notions of vulnerability, which it applies differently to refer to both subjects and information systems. The elderly person, who is not inherently vulnerable, but may potentially be more so, finds specific consideration in the legal dimension of digital. In this sense can be understood the articulated normative response to the problem of algorithmic discrimination. The provisions of new technology law — public and private — thus work together to offer protection to the subject in his or her specific personal circumstances.

Nella molteplicità delle sfide poste dall'innovazione tecnologica, la centralità del concetto di sicurezza emerge sul piano giuridico e acquista nuova rilevanza attraverso le declinazioni della c.d. cibersicurezza. In questo ambito il linguaggio del diritto accoglie diverse nozioni di vulnerabilità, che applica differentemente per riferirsi tanto ai soggetti quanto ai sistemi informatici. La persona anziana, che non è per ciò solo vulnerabile, ma che potenzialmente può esserlo in misura maggiore, trova una specifica considerazione nella dimensione giuridica del digitale. In tal senso può intendersi l'articolata riposta normativa al problema della discriminazione algoritmica. Le disposizioni

* Research Fellow in Private Law at the University of Padova.

§ This contribution develops the presentation at the conference 'Cybersecurity e prodotti: sfide e responsabilità nelle società interconnesse', 29 November 2024, Italian Supreme Court. The article is a product of research conducted within the research programme Age-it, Ageing well in an ageing society, Spoke 5, Care sustainability in an ageing society, of the National Recovery and Resilience Plan.

proprie del diritto delle nuove tecnologie – pubblico e privato – operano quindi congiuntamente per offrire tutela al soggetto, nelle sue specifiche condizioni personali.

1. Law and technological innovation. Vulnerability of technology and vulnerability of the subject

Technology transcends the boundaries of science, draws on the most diverse contexts, and evolves at an increasingly rapid pace. The acceleration of innovation is also determined by the impetus provided by the economy and the market[1]. There is no aspect of human life that is not now pervaded by technology[2] and legal studies have long since become aware of the close connection between technology and law[3].

The relevance of the topic is transversal, institutional in nature and questions the foundations of the legal system and the values of the order[4]. Its constitutional scope can be observed by looking at the impact of the newest technologies on the guarantee of personal rights[5].

Private law is also called upon to respond to the questions posed by technological innovation[6]. Understanding whether and how traditional categories can be preserved in the face of the sometimes sudden developments in technology is a task for the interpretation of law[7].

[1] G. Alpa, *Il mercato unico digitale*, in «Contratto e impresa Europa», I (2021), pp. 1 ff.

[2] P. Perlingieri, *Note sul «potenziamento cognitivo»*, in «Tecnologie e diritto», I (2021), pp. 209 ff.; Id., *Sul trattamento algoritmico dei dati*, in «Tecnologie e diritto», (2020), pp. 181 ff.

[3] N. Irti, *Il diritto nell'età della tecnica*, Editoriale Scientifica, Napoli 2007. Cf. G. Di Rosa, *Profili giuridici dell'esistenza,* Giappichelli, Torino, 2022, pp. 153 ff.; E. Maestri, *Lex informatica. Diritto, persona e potere nell'età del cyberspazio,* Edizioni Scientifiche Italiane, Napoli, 2015.

[4] G. Alpa (ed.), *Diritto e intelligenza artificiale*, Pacini, Pisa 2020. Cf. E. Battelli, *Necessità di un umanesimo tecnologico: sistemi di intelligenza artificiale e diritti della persona,* in «Il diritto di famiglia e delle persone», III (2022), pp. 1096 ff.

[5] Cf. C. Casonato, M. Fasan, S. Penasa (eds.), *Diritto e intelligenza artificiale, monographic section* in «DPCE online», I (2022), pp. 155 ff.; A. D'Aloia (ed.), *Intelligenza artificiale e diritto. Come regolare un mondo nuovo,* FrancoAngeli, Milano, 2021, e, *ivi,* Id., *Il diritto verso "il mondo nuovo". Le sfide dell'Intelligenza Artificiale,* pp. 7 ff.; Id., *Il diritto e l'incerto mestiere del vivere. Ricerche di biodiritto,* Cedam, Padova, 2021, pp. 203 ff.; Id., *Il diritto verso "il mondo nuovo". Le sfide dell'Intelligenza Artificiale,* in «BioLaw Journal-Rivista di BioDiritto», I (2019), pp. 3 ff.

[6] C. Perlingieri, *Diritto privato delle nuove tecnologie: contenuti e competenze,* in «Tecnologie e diritto», II (2021), pp. 70 ff. See also A. Alpini, *La trasformazione digitale nella formazione del civilista, ivi*, pp. 1 ff.

[7] R. Clarizia, *Internet: gli interrogativi del civilista,* in «Diritto di Internet», III (2022), pp. 3 ff. Cf. M. Franzoni, *Il digitale, la rete, l'IA e la responsabilità civile,* in «Jus civile», II (2024), pp. 208 ff.

In the context of the digitisation process[8], which pervades all aspects of contemporary society, the concept of digital governance emerges as a pivotal consideration. European Union law assumes a prominent role in this context[9].

In 2020, with its Communication Shaping Europe's digital future, the European Commission identified three key objectives in the pursuit of digitalisation for the next five years: technology at the service of people, a fair and competitive economy and an open, democratic and sustainable society. The stages of this journey were then defined in the Communication of 9 March 2021, Digital Compass 2030, whose 'digital targets' were identified in Decision No. 2481 of 14 December 2022 of the European Parliament and of the Council. The 'European Declaration on Digital Rights and Principles for the Digital Decade' of 15 December 2022, solemnly proclaimed by the European Parliament, the Council and the Commission, enshrines the official recognition of the rights of the individual and the principles of the European Union in the digital transformation. Artificial Intelligence is also part of the European digitisation process, which the aforementioned acts consider in various ways, in programmatic, promotional and guiding terms.

The landscape of EU law in this area does not only consist of soft law, but also includes hard law. The most recent regulations include: the General Data Protection Regulation, EU reg. 2016/679 (GDPR); the Cybersecurity Act, EU reg. 2019/881; the Data Governance Act, EU reg. 2022/868; the Digital Markets Act, EU reg. 2022/1925; the Digital Services Act, EU reg. 2022/2065; the Data Act, EU reg. 2023/2854; the Artificial Intelligence Act, EU reg. 2024/1689 (AI Act); the Cyber Resilience Act, EU reg. 2024/2847; the Cyber Solidarity Act, EU reg. 2025/38; the European Health Data Space Regulation, EU reg. 2025/327.

This is an ever-changing regulatory framework. The proposed Artificial Intelligence Liability Directive (AILD), put forward in 2022 and part of the regulatory 'package' dedicated to Artificial Intelligence (which included, in

[8] T.E. Frosini, *L'ordine giuridico del digitale. The legal order of digital,* in «Giurisprudenza costituzionale», I (2023), pp. 377 ff.; A. Alpini, *La trasformazione digitale nella formazione del civilista,* cit., pp. 1 ff.; A. Sandulli, *Lo «Stato digitale». Pubblico e privato nelle infrastrutture digitali nazionali strategiche,* in «Rivista trimestrale di diritto pubblico», II (2021), pp. 513 ff. See also R. Cavallo Perin (ed.), *L'amministrazione pubblica con i big data: da Torino un dibattito sull'intelligenza artificiale,* Università degli Studi di Torino, Torino, 2021. Cf. G. Pisani, *Piattaforme digitali e autodeterminazione. Relazioni sociali, lavoro e diritti al tempo della "governamentalità algoritmica",* Mucchi, Modena, 2023.

[9] C. Camardi, *Sulla Governance digitale europea: una proposta di confronto,* in «Accademia», I (2023), pp. 7 ff.; A. Alpini, *La sovranità digitale europea,* in «TIGOR», (2022), pp. 144 ff. See also S. Corso, *Trasformazione digitale, digitalizzazione della sanità e intelligenza artificiale. La Dichiarazione europea sui diritti e i principi digitali e il programma strategico per il decennio digitale 2030,* in *federalismi.it, Osservatorio di diritto sanitario*, (2023).

addition to the AI Act, the Product Liability Directive (PLD), repealing Directive 85/374/EEC), was withdrawn in February 2025[10].

On 17 May 2024, the Framework Convention on Artificial Intelligence, Human Rights, Democracy and the Rule of Law was adopted by the Council of Europe. On 5 September 2024, the European Commission signed this first international legally binding instrument on Artificial Intelligence for the Union, which is compatible with EU law and the AI Act. The intertwining of human rights profiles, democracy and the rule of law in the regulation of the technological phenomenon runs through the text of the Convention, highlighting the interconnection between the level of the individual and that of the community, between the private and the public.

Between the lines of this growing and heterogeneous body of law, among the many legal concepts, one can discern that of vulnerability.

The term 'vulnerability' is employed in the context of these provisions with regard to both the subject and the machine. In the latter sense, for instance, the AI Act refers to the vulnerability of the system or the vulnerability of the AI in recitals 76, 77 and 110 and in Article 15, with a view to regulating particular aspects in order to ensure the Cybersecurity of the AI system. Similarly, it can be noted that this word is used widely to refer to the characteristics of technology in the Cybersecurity Act, and even more so in the Implementing Regulation No. 482 of 2024, which devotes an entire chapter, Chapter VI, to this concept, or in the Cyber Resilience Act. The provisions that address technological vulnerability dictate rules aimed at reducing the risks associated with the technology itself, in other words, promoting security.

In the different meaning concerning the natural person, the term vulnerability is used in the legislation to create a discipline that ensures respect for the fundamental freedoms and rights of each individual, specifically taking into account the specific conditions of individuals or groups of individuals[11]. There is evidence of greater maturity in the legal texts on this point, which are increasingly addressing these aspects. Thus, while in the GDPR the notion surfaced timidly only in a recital, no. 75, to highlight how potentially harmful processing of data of 'vulnerable natural persons' may pose risks to people's rights and freedoms[12], in the text of the AI Act the idea belongs in several conceptual strands, in several articles and recitals[13].

[10] Communication from the Commission, *Moving forward together: A Bolder, Simpler, Faster Union,* COM(2025) 45 final, 11 february 2025.

[11] Cf. G. Malgieri, *Vulnerability and Data Protection Law,* Oxford University Press, Oxford, 2023.

[12] C. Irti, *La persona anziana tra famiglia e società. Problematiche di diritto civile,* Pacini, Pisa, 2023, p. 106.

[13] Thus, expressly Articles 5, 7, 9, 60, 79 and 95; recitals 29, 48, 58, 60, 67, 93, 110, 132, 141 and 165.

The individual may find himself in a situation of vulnerability[14] for various reasons, social, economic or health-related, or for any other personal condition that exposes him to harm, claiming particular protection[15]. Several potential vulnerability factors are linked to a specific situation, which is not in itself a condition of vulnerability: elderly age[16].

2. The elderly and the digital

From the simplest everyday tasks to more complex ones, such as estate or succession planning[17], or health-related[18], technological innovation can certainly meet the needs of elderly age. But it is also certain that it is accompanied by new risks for the individual and his rights.

It is evident that the population of EU Member States is undergoing a process of progressive ageing[19], a phenomenon that is particularly

[14] On vulnerability in legal theory, ex multis, B. Pastore, *Semantica della vulnerabilità, soggetto, cultura giuridica,* Giappichelli, Torino 2021; E. Pariotti, *Vulnerabilità ontologica e linguaggio dei diritti,* in «Ars interpretandi», II (2019), pp. 155 ff.; G. Zanetti, *Filosofia della vulnerabilità. Percezione, discriminazione, diritto,* Carocci, Roma, 2019; O. Giolo, B. Pastore (eds.), *Vulnerabilità. Analisi multidisciplinare di un concetto*, Carocci, Roma, 2018; T. Casadei (ed.), *Diritti umani e soggetti vulnerabili. Violazioni, trasformazioni, aporie*, Giappichelli, Torino, 2012.

[15] An idea of vulnerability with a legal value relates to the situation in which a human being finds himself in need of protection. S. Giova, *Minori in condizione di particolare vulnerabilità,* in «Rivista giuridica del Molise e del Sannio», III (2022), p. 168. Cf. A. Gentili, *La vulnerabilità sociale. Un modello teorico per il trattamento legale,* in «Rivista critica di diritto privato», I (2019), pp. 42 ff.; N. Lipari, *Vulnerabilità esistenziale e strumenti di tutela*, in «Annali della SISDiC», II (2018), pp. 3 ff.

[16] P. Perlingieri, *Diritti della persona anziana, diritto civile e stato sociale,* in «Rassegna di diritto civile», I (1990), pp. 80 ff., now in Id., *La persona e i suoi diritti,* Esi, Napoli, 2005, pp. 345 ff. Cf. L. Mengoni, *La tutela giuridica della vita materiale nelle varie età dell'uomo,* in «Rivista trimestrale di diritto e procedura civile», IV (1982), pp. 1117 ff. and in Id., *Diritto e valori,* il Mulino, Bologna, 1985, pp. 123 ff.; P. Stanzione (ed.), *Anziani e tutele giuridiche,* Esi, Napoli, 1991; Id., «Anziani (assistenza agli)», in *Enc. del dir.*, Agg. III, Giuffrè, Milano, 1999, pp. 112 ff.; C.M. Bianca, *Senectus ipsa morbus?*, in «Rassegna di diritto civile», II (1998), pp. 241 ff.; M. Tamponi, *Nel diritto della terza età. Le rughe tra giudizio e pregiudizio,* Rubbettino, Soveria Mannelli, 2021; M. Cascione, *Il lato grigio del diritto. Invecchiamento della popolazione e tutela degli anziani in prospettiva comparatistica,* Giappichelli, Torino, 2022; C. Irti, *La persona anziana tra famiglia e società. Problematiche di diritto civile,* cit.; Ead., *La persona anziana: sulla rilevanza giuridica della nozione,* in AA.VV., *Liber Amicorum per Paolo Zatti*, Jovene, Napoli, 2023, pp. 1215 ff.; Ar. Fusaro, M. Piccinni (eds.), *Un diritto gentile per la persona anziana*, Pacini, Pisa, 2024.

[17] One thinks of the use of technologies that enable the infirm or disabled person to manifest his or her will, such as, for instance, when drafting a will. See A.A. Mollo, *Il testatore vulnerabile e le esigenze di rinnovamento delle norme sulla successione a causa di morte a tutela dei diritti delle persone con disabilità: il testamento digitale videoregistrato,* in «Familia», V (2023), pp. 671 ff.

[18] Cf. G. Capilli, *Diritto privato sanitario. Fondamenti,* 2ª ed., Pacini, Pisa, 2025.

[19] The number of people aged sixty-five and over, including those in their eighties, is growing, while the number of young people under the age of twenty is falling, with an increase in the average age. Statistical survey data can be found in *www.istat.it.*

pronounced in Italy. Concurrently, a decline in the birth rate is being observed[20] in a context of demographic crisis and, above all, thanks to scientific and technological progress, an increase in well-being and survival, also thanks to new possibilities for treatment and cures. In this context, new technologies have emerged as a pivotal factor[21]. The enhanced longevity that has been achieved is, to a considerable extent, attributable to these technological advancements, which have also facilitated the identification of solutions to the challenges posed by advanced age.

A key objective of EU digitisation legislation is to ensure the effective protection of individuals' rights in the context of technological development. In pursuit of this objective, particular consideration is given to the position of vulnerable individuals, including the elderly. This commitment to inclusivity is formally articulated in the European Declaration on Digital Rights and Principles, which reinforces solidarity and the principle of egalitarian inclusion. In so doing, it is asserted that a «digital transformation that leaves nobody behind' should be pursued, a commitment which includes the elderly»[22]. With regard to participation in the digital public space, after recognising the importance of a trustworthy digital environment that everyone can access and in which democratic debate can take place, it calls on institutions to «limit the exploitation of vulnerabilities and biases, namely through targeted advertising»[23]. In this context, therefore, digital technology appears to be a tool to ensure inclusion, sustainability and support for the elderly, in harmony with the principles of the European Treaties[24]. «*Technology should be used to unite, and not divide, people. The digital transformation should contribute to a fair and inclusive society and economy in the EU*»[25].

In the Italian legal system, legislative decree no. 29 of 15 March 2024, which implements the delegation referred to in Articles 3, 4 and 5 of law no.

[20] G. Alpa, *La famiglia nell'età postmoderna. Nuove regole, nuove questioni,* in AA.VV., *Liber Amicorum per Paolo Zatti,* Jovene, Napoli, 2023, pp. 1435 ff., and in «Rivista trimestrale di diritto e procedura civile», II (2023), pp. 363 ff., spec. pp. 372 f.

[21] Pioneering studies include S. Rodotà, *Tecnologie e diritti,* il Mulino, Bologna, 1995. See P. Perlingieri, S. Giova, I. Prisco (eds.), *Il trattamento algoritmico dei dati tra etica, diritto ed economia,* Esi, Napoli, 2020. Cf. E. Battelli (ed.), *Diritto privato digitale*, 2ª ed., Giappichelli, Torino, 2024; C. Perlingieri, I. Martone (eds.), *Nuove tecnologie e cultura del diritto civile*, Esi, Napoli 2023.

[22] «It should benefit everyone, achieve gender balance, and include notably elderly people, people living in rural areas, persons with disabilities, or marginalised, vulnerable or disenfranchised people and those who act on their behalf. It should also promote cultural and linguistic diversity». Chapter II, lett. *b.*

[23] Chapter IV, lett. *f.*

[24] Cf. A. Alpini, *Sull'approccio umano-centrico all'intelligenza artificiale. Riflessioni a margine del "Progetto europeo di orientamenti etici per una IA affidabile",* in *www.comparazionedirittocivile.it,* (2019).

[25] Chapter II, 2, of the European Declaration on Digital Rights and Principles for the Digital Decade.

33 of 23 March 2023, lays down provisions on policies in favour of the elderly. 'Elderly person' is defined by Article 2(1) as one who has reached the age of sixty-five and 'grand elderly person' eighty years old[26]. The aim, as expressed in Article 1, is to promote dignity and autonomy, social inclusion, active ageing and the prevention of frailty in the elderly population.

In the direction indicated by the Union law, Law 33/2023 has been oriented, including in article 3, paragraph 2, letter a, no. 7, among the guiding principles and criteria for the delegation of powers on the subject of active ageing, the promotion of social inclusion and the prevention of frailty, with regard to interventions for active ageing and the promotion of the autonomy of the elderly, the valorisation of computer literacy actions and practices that enable the use of new technologies suitable for promoting the knowledge and civil and social participation of the elderly, supporting autonomy in the management of their lives and guaranteeing full access to services and information.

Legislative Decree 29/2024 then dictated measures to support the elderly population with regard to computer literacy and digital facilitation. Thus, Article 19 provides for the promotion of activities to train the digital skills of the elderly and support them in using the services provided online by public administrations, implementing the provisions of the PNRR. Similarly, in order to reduce the digital divide between generations[27], Article 20 allows second cycle education and training institutions, within their autonomy, to encourage the creation of training pathways that promote the acquisition of knowledge and skills in the use of digital tools by elderly people, also involving students.

Digitalisation or, more generally, new technologies relate to the elderly, offering new opportunities, support and assistance of various kinds. They can improve the quality of life, promote independence, create safe environments, encourage socialisation, provide mobility solutions or offer opportunities for learning and entertainment. In this sense, the relevant legislation promotes the use of technology while guaranteeing the rights of the individual. In this way, it pursues aims of safety and risk prevention and management in a general way, so that they can be applied to a wide variety of situations, including those specific to an elderly person.

[26] These are the definitions laid down in letters a and b of Article 2(1). In letters c and d of the same subsection the terms "dependent elderly person" and "specific care need of the dependent elderly person" are defined. Critically, on the Act, see F. Cembrani, D. De Leo, *La legge delega in materia di politiche in favore degli anziani: le insidie delle molte deleghe in bianco lasciate al Governo e di una elder law italiana che sembra aver messo nell'angolo la componente di cura della non autosufficienza,* in «Rivista italiana di medicina legale», (2023), pp. 447 ff.

[27] Cf. A. Alpini, *La trasformazione digitale nella formazione del civilista,* cit.

3. Safety, security and risk. A specific problem: algorithmic discrimination

Ensuring safety, security and managing risk are objectives peculiar to regulation of newer technologies, but they are not new objectives. Consumer law has long been aware of the implications of these issues in relation to the concept of product and in the link between safety and security. Moreover, in recent years, under the impetus of European Union law, it has been constantly updated precisely to take account of technological developments. Think of Legislative Decree No. 170 of 4 November 2021, implementing Directive No. 771 of 2019 on certain aspects of contracts for the sale of goods, or Legislative Decree No. 173 of 4 November 2021, implementing Directive No. 770 of 2019 on certain aspects of contracts for the supply of digital content and digital services[28]. Consider the impact of the application of EU reg. 2023/988, on general product safety[29], or EU Reg. 2024/2847, on cyber resilience, and the adjustments that may be required to internal discipline. Think also of the changes that will be necessary in the area of producer liability, also due to the provisions on Artificial Intelligence.

Looking at the use of Artificial Intelligence systems and the risks associated with them, we can see that so-called algorithmic discrimination is a specific issue that affects vulnerable people[30].

The algorithms underlying Artificial Intelligence systems operate on the basis of their data. If the machine learns from a data set in which one or more groups of people are excluded or under-represented, decisions based on the output thus provided may prejudice the person belonging to one of those groups[31].

[28] G. De Cristofaro, *Legislazione italiana e contratti dei consumatori nel 2022: l'anno della svolta. Verso un diritto "pubblico" dei (contratti dei) consumatori?,* in «Nuove leggi civili commentate», I (2022), pp. 1 ff.; A. De Franceschi, *Italian Consumer Law after the Transposition of Directives (EU) 2019/770 and 2019/771*, in «Journal of European Consumer and Market Law», II (2022), pp. 72 ff.

[29] Cf. N. Chiricallo, *Il "richiamo" dei prodotti pericolosi e i nuovi rimedi civilistici esperibili dall'acquirente-consumatore. L'art. 37 del Regolamento UE 2023/988 sulla sicurezza generale dei prodotti,* in «Nuove leggi civili commentate», III (2024), pp. 669 ff.

[30] A.G. Grasso, *The Bad Algorithm. Automated Discriminatory Decisions in the European General Data Protection Regulation,* Esi, Napoli, 2022; E. Falletti, *Discriminazione algoritmica. Una prospettiva comparata,* Giappichelli, Torino, 2022. Cf. B. Parenzo, *La profilazione algoritmica nel prisma dell'autonomia privata,* Esi, Napoli, 2024; Id., *Profilazione e discriminazione. Dal GDPR alla Proposta di Regolamento sull'IA,* in «Tecnologie e diritto», I (2023), pp. 105 ff.; Id., *Profilazione e discriminazione. Dal GDPR alla proposta di Regolamento sull'intelligenza artificiale,* in *La via europea per l'intelligenza artificiale*, edited by C. Camardi, *La via europea per l'intelligenza artificiale,* Cedam, Padova, 2022, pp. 335 ff. See also the focus by O. Giolo, *L'algoritmo alla prova del caso concreto: stereotipi, serializzazione, discriminazione*, in «GenIUS», II (2022).

[31] The principle of non-discrimination, together with the principle of knowability of the decision-making process and non-exclusivity of the decision, has been identified by administra-

It is important to consider the potential applications of Artificial Intelligence in healthcare, particularly when it comes to the assessment of health conditions[32]. The use of such devices necessitates the consideration of the representativeness of the data set, which is crucial for the accuracy of the assessment. In the event that the data set used by the device does not adequately represent the characteristics of the elderly individual, the assessment may not be reliable.

In the explanatory memorandum accompanying the proposal for a regulation on Artificial Intelligence, the Commission pointed out that the envisaged framework supplemented EU law on non-discrimination with specific requirements aimed at minimising the risk of algorithmic discrimination, in particular in relation to the design and quality of the datasets used for the development of AI systems. Moreover, the Council of Europe Convention on Artificial Intelligence also expressly refers to the risk of discrimination inherent in this technology and the need for protection of the vulnerable[33].

The solution to this problem offered by the EU landscape rests on the European Data Strategy. As stated in recital 68 of the AI Act, «European common data spaces established by the Commission and the facilitation of data sharing between businesses and with government in the public interest will be instrumental to provide trustful, accountable and non-discriminatory access to high-quality data for the training, validation and testing of AI systems. For example, in health, the European health data space will facilitate non-discriminatory access to health data and the training of AI algorithms on those data sets, in a privacy-preserving, secure, timely, transparent and trustworthy manner, and with an appropriate institutional governance. Relevant competent authorities, including sectoral ones, providing or supporting the access to data may also support the provision of high-quality data for the training, validation and testing of AI systems»[34]. The European Health Data Space, established by EU reg. 2025/327[35], will enable people to control and use

tive jurisprudence as the rule underlying the operation of the algorithm. Cons. Stato, 13.12.2019, no. 8472, in «Nuova giurisprudenza civile commentata», 2020, I, pp. 809 ff., with a note from R. Mattera, *Decisioni algoritmiche. Il Consiglio di Stato fissa i limiti.*

[32] See S. Corso, *L'intelligenza artificiale e il trattamento dei dati relativi alla salute,* in «Responsabilità medica», (2023), pp. 317 ff.

[33] The preamble to the Convention explicitly refers to the «*risks of discrimination in digital contexts, particularly those involving artificial intelligence systems, and their potential effect of creating or aggravating inequalities, including those experienced by women and individuals in vulnerable situations, regarding the enjoyment of their human rights and their full, equal and effective participation in economic, social, cultural and political affairs*». See also Article 18.

[34] This is the recital corresponding to No 45 of the 2021 proposal.

[35] S. Corso, *Autodeterminazione e dati sanitari*, Giappichelli, Torino, 2025; Id., *Lo spazio europeo dei dati sanitari. Prime riflessioni sul Regolamento UE 2025/327,* in «Nuove leggi civ.

their own electronic health data both in their own country and in other Member States, promote a single market for digital health services and products, and provide a coherent, reliable and efficient framework for the use of these data in research, innovation, policy-making and regulation, while respecting the Union's high standards of data protection. The Data Act and the Data Governance Act help facilitate the sharing and use of data.

Therefore, since the use of Artificial Intelligence may adversely affect the rights enshrined in the Charter of Fundamental Rights of the European Union, the provisions of the AI Act aim to ensure a high level of protection of fundamental rights and to address various sources of risk through a risk-based approach, as announced by the Commission in the explanatory memorandum to the proposal. The setting of the AI Act is thus characterised by the correlation between Artificial Intelligence and product[36] and by the centrality of the notion of risk. A distinction is made between low — or minimal — risk systems and high-risk systems, for which ad hoc rules are dictated, and those whose risk is unacceptable, for which special prohibitions apply.

The AI Act stipulates, in Article 10, for systems classifiable as high-risk that use techniques involving the use of data to train Artificial Intelligence models, that these must be developed on the basis of training, validation and testing datasets that meet quality criteria[37]. A comparison of the 2021 proposal[38] and the 2024 approved text reveals a shift towards a more pronounced emphasis on combating discrimination. Consequently, the criteria, in addition to those aimed at ensuring the correctness and adequacy of data sets, encompass the assessment of «possible biases that are likely to affect the health and safety of

comm.», I (2025), pp. 563 ff. With regard to the 2022 proposal, see Id., *Lo spazio europeo dei dati sanitari: la Commissione Europea presenta la proposta di regolamento,* in *www.federalismi.it*, Osservatorio di diritto sanitario, (2022); Id., *European Health Data Space. La Commissione europea presenta la proposta di Regolamento sullo spazio europeo dei dati sanitari,* in *www.rivistaresponsabilitamedica.it,* (2022); Id., *Una strategia europea per i dati, anche sanitari*, *ivi,* (2021).

[36] In the interconnection between Artificial Intelligence and responsibility. G. Capilli, *Responsabilità e robot*, in «Nuova giurisprudenza civile commentata», II (2019) pp. 621 ff.; Ar. Fusaro, *Quale modello di responsabilità per la robotica avanzata? Riflessioni a margine del percorso europeo*, in «Nuova giurisprudenza civile commentata», II (2020), pp. 1344 ff.; G. Finocchiaro, *Intelligenza artificiale e responsabilità,* in «Contratto e impresa», II (2020), pp. 713 ff.; U. Salanitro, *Intelligenza artificiale e responsabilità: la strategia della commissione europea,* in «Rivista di diritto civile», VI (2020), pp. 1246 ff. Cf. C. Perlingieri, *L'incidenza dell'utilizzazione della tecnologia robotica nei rapporti civilistici,* in «Rassegna di diritto civile», (2015), pp. 1235 ff.; U. Pagallo, *The Laws of Robots. Crimes, Contracts, and Torts*, Springer, Berlino, 2013, pp. 152 ff.

[37] C. Irti, *L'uso delle "tecnologie mobili" applicate alla salute: riflessioni al confine tra la forza del progresso e la vulnerabilità del soggetto anziano,* in «Persona e mercato», I (2023), p. 45. See in this respect recital 67 of the AI Act and recital 44 of the draft Regulation.

[38] Cf. E. Battelli, *Necessità di un umanesimo tecnologico: sistemi di intelligenza artificiale e diritti della persona,* cit., pp. 1113 ff.

persons, have a negative impact on fundamental rights or lead to discrimination prohibited under Union law, especially where data outputs influence inputs for future operation» as well as the appropriate measures to detect, prevent and mitigate the possible distortions thus identified[39].

Article 6 of the AI Act, by introducing the criteria for classifying an AI system as high-risk, now circumscribes its definitional area and, in particular, in para. 3, excludes Annex III systems, which, according to para. 2, should be considered high-risk, when they do not present «significant risk of harm to the health, safety or fundamental rights of natural persons, including by not materially influencing the outcome of decision making». It is the provider, as stated in para. 4, who will assess whether or not the system is high risk.

In the area of high-risk systems, Artificial Intelligence systems used in healthcare are included if they are products or safety components of a product belonging to the category of 'medical device'[40]. Software or technological tools that make use of algorithms for the health of the elderly may or may not belong there. Apps that can be downloaded online and have no medical destination, for instance, are not likely to be included[41].

However, there are various factors which, if not adequately considered by the Artificial Intelligence system, can compromise the use of algorithms, and therefore it is not only age that matters. Each factor can then be combined with others, resulting in double or multiple algorithmic discrimination. For example, imagine a system that takes age into account, but in which homosexual or black people are not included, or specifically homosexual elderly people or black elderly people. Or, again, consider a system that does not include elderly women. Algorithmic discrimination may thus also be double or multiple discrimination. Algorithmic bias compromises the inclusiveness of technology[42].

An information and communication technology product that determines or does not prevent algorithmic discrimination cannot be said to be safe. Or perhaps Cyber-secure[43].

[39] Article 10(2), lett. f and g. Cf. the different wording of recital 67 of the text approved in May 2024, compared to the corresponding recital 44 of the 2021 proposal.

[40] By virtue of paragraph 1's reference to Annex I, which expressly refers to Regulation No. 745 of 2017.

[41] Cf. C. Irti, *L'uso delle "tecnologie mobili" applicate alla salute: riflessioni al confine tra la forza del progresso e la vulnerabilità del soggetto anziano*, cit., p. 35.

[42] With regard to the use of Artificial Intelligence for women's health, see S. Corso, *Sulla female technology o Femtech. Intelligenza artificiale, salute della donna e volontà della persona*, in «BioLaw Journal-Rivista di BioDiritto», III (2023), pp. 81 ff.

[43] The Cybersecurity Act defines 'Cybersecurity' in Article 2, no. 1, as «the activities necessary to protect network and information systems, the users of such systems, and other persons affected by cyber threats», and, in Article 2, no. 8, 'cyber threat' as «any potential circumstance, event or action that could damage, disrupt or otherwise adversely impact network

4. Product Cybersecurity and protection of the individual

With Legislative Decree No. 138 of 4 September 2024 (so-called NIS Legislative Decree), the Italian legislature transposed EU Directive No. 2555 of 2022 (NIS 2), concerning measures for a high common level of Cybersecurity in the Union[44]. Several provisions explicitly refer to information and communication technology (ICT) products. Also in this context, the term 'vulnerability' relates to the machine. As delineated in Article 2(ii), a vulnerability is defined as a weakness, susceptibility or defect in ICT products or ICT services that can be exploited by a cyber threat. This constitutes a pivotal concept within the Cybersecurity framework. To ensure the reduction of risk, the minimisation of damage, and accountability and transparency in the management of system vulnerabilities, a coordinated disclosure procedure is outlined in Article 16. This procedure facilitates the reporting and resolution of issues that are intrinsically linked to the product.

Vulnerability management, including the promotion and facilitation of coordinated disclosure of vulnerabilities as referred to in Article 16 is one of the strategic measures that Article 9(3) includes in the 'national Cybersecurity strategy'.

The notion of Cybersecurity takes on a very broad scope and encompasses the legal system as a whole. From the public dimension, the Cybersecurity regime now ranges as far as private law categories[45].

Faced with the magnitude of the technological phenomenon and the multiplicity of the needs of individuals, each in his or her own concrete circumstances, the boundaries between public and private law seem to blur here. The law of the new technologies, whether public or private, takes into account the particularity of the subject – elderly, fragile, vulnerable – in order to guarantee protection in existential legal situations[46].

and information systems, the users of such systems and other persons». For a study on cyber risk, see A. Candian, G. Romagnoli (eds.), *Cyber risk e la postmodernità cibernetica,* Esi, Napoli, 2024.

[44] Cf. F. Casarosa, G. Comandè, *Aspettando la NIS2: ovvero il diritto privato della cybersecurity*, in «Il diritto dell'informazione e dell'informatica», I (2024), pp. 29 ff.; F.P. Micozzi, *Sicurezza informatica. Obblighi e responsabilità dopo il recepimento della NIS2 e la L. 90/2024,* Cedam, Padova, 2024.

[45] F. Casarosa, G. Comandè, cit.

[46] P. Perlingieri, *Il diritto civile nella legalità costituzionale secondo il sistema italo-europeo delle fonti*, III, *Situazioni soggettive*, 4ª ed., Esi, Napoli, 2020, pp. 1 ff. Cf. S. Giova, *La persona nel contesto digitale: considerazioni conclusive*, in «Tecnologie e diritto», II (2023), pp. 326 ff. As stated in point 1 of the European Declaration on Digital Rights and Principles for the Digital Decade, Chapter I, 'Putting People at the Centre of Digital Transformation': «People are at the centre of the digital transformation in the European Union. Technology should serve and benefit all people living in the EU and empower them to pursue their aspirations, in full security and respect for their fundamental rights».

The concept of technological vulnerability also contributes, at least indirectly, by participating in the guarantee of security and the prevention and management of risk, to the more general function of protection of the person and, in this, the consideration of the elderly person also finds its place. The set of new rights and obligations recognised for those operating in the technology market or more broadly in the context of technology relates to the traditional institutions of the legal system to protect the individual[47]. The new Cybersecurity provisions, as a whole, also contribute to the protection of the individual, including the vulnerable, such as the elderly person.

Through constitutional principles, primarily the personalist principle[48], a concept of vulnerability[49] is transmitted to the whole of law – and also to civil law – which, in various ways, but always, draws on the semantics of dignity[50]. Thus, from the legal order of the market to the legal order of technology[51], the role of private law appears essential.

[47] E. Battelli, *Necessità di un umanesimo tecnologico: sistemi di intelligenza artificiale e diritti della persona,* cit., pp. 1118 ff.

[48] P. Perlingieri, *Diritti della persona anziana, diritto civile e stato sociale*, cit. Cf. Id., *Il diritto civile nella legalità costituzionale secondo il sistema italo-europeo delle fonti, III, Situazioni soggettive,* cit., pp. 18 ff.

[49] Ar. Fusaro, *L'atto patrimoniale della persona vulnerabile*, Jovene, Napoli, 2019.

[50] P. Zatti, *Note sulla semantica della dignità, in ID., Maschere del diritto, volti della vita,* Giuffrè, Milano, 2009, pp. 29 ff.

[51] N. Irti, *L'ordine giuridico del mercato*, Laterza, Roma-Bari 1998; L. d'Avack, *Ordine giuridico e ordine tecnologico,* 2ª ed., Giappichelli, Torino, 1998. Cf. T.E. Frosini, *L'Ordine giuridico del digitale*, in *Strategia dei dati e intelligenza artificiale. Verso un nuovo ordine giuridico del mercato,* by V. Falce (ed.), Giappichelli, Torino, 2023, pp. 17 ff.

BIBLIOGRAPHY

Alpa G., *La famiglia nell'età postmoderna. Nuove regole, nuove questioni*, in Aa.Vv., *Liber Amicorum per Paolo Zatti*, Jovene, Napoli 2023, and in «Rivista trimestrale di diritto e procedura civile», II (2023).

Alpa G., *Il mercato unico digitale*, in «Contratto e impresa Europa», I (2021).

Alpa G. (ed.), *Diritto e intelligenza artificiale*, Pacini, Pisa, 2020.

Alpini A., *La sovranità digitale europea*, in «TIGOR», (2022).

Alpini A., *La trasformazione digitale nella formazione del civilista*, in «Tecnologie e diritto», II (2021).

Alpini A., *Sull'approccio umano-centrico all'intelligenza artificiale. Riflessioni a margine del "Progetto europeo di orientamenti etici per una IA affidabile"*, in *www.comparazionedirittocivile.it*, (2019).

Battelli E. (ed.), *Diritto privato digitale*, 2ª ed., Giappichelli, Torino, 2024.

Battelli E., *Necessità di un umanesimo tecnologico: sistemi di intelligenza artificiale e diritti della persona*, in «Il diritto di famiglia e delle persone», III (2022).

Bianca C.M., *Senectus ipsa morbus?*, in «Rassegna di diritto civile», II (1998).

Camardi C., *Sulla Governance digitale europea: una proposta di confronto*, in «Accademia», I (2023).

Candian A., Romagnoli G. (eds.), Cyber risk *e la postmodernità cibernetica*, Esi, Napoli, 2024.

Capilli G., *Diritto privato sanitario. Fondamenti*, 2ª ed., Pacini, Pisa, 2025.

Capilli G., *Responsabilità e robot*, in «Nuova giurisprudenza civile commentata», II (2019).

Casadei T. (ed.), *Diritti umani e soggetti vulnerabili. Violazioni, trasformazioni, aporie*, Giappichelli, Torino, 2012.

Casarosa F., Comandè G., *Aspettando la NIS2: ovvero il diritto privato della* cybersecurity, in «Il diritto dell'informazione e dell'informatica», I (2024).

Cascione M., *Il lato grigio del diritto. Invecchiamento della popolazione e tutela degli anziani in prospettiva comparatistica*, Giappichelli, Torino, 2022.

Casonato C., Fasan M., Penasa S. (eds.), *Diritto e intelligenza artificiale*, monographic section in «DPCE online», I (2022).

Cavallo Perin R. (ed.), *L'amministrazione pubblica con i* big data*: da Torino un dibattito sull'intelligenza artificiale*, Università degli Studi di Torino, Torino, 2021.

Cembrani F., De Leo D., *La legge delega in materia di politiche in favore degli anziani: le insidie delle molte deleghe in bianco lasciate al Governo e di una* elder law *italiana che sembra aver messo nell'angolo la componente di cura della non autosufficienza*, in «Rivista italiana di medicina legale», (2023).

Chiricallo N., *Il "richiamo" dei prodotti pericolosi e i nuovi rimedi civilistici esperibili dall'acqui-*

rente-consumatore. L'art. 37 del Regolamento UE 2023/988 sulla sicurezza generale dei prodotti, in «Nuove leggi civili commentate», III (2024).

Clarizia R., *Internet: gli interrogativi del civilista*, in «Diritto di Internet», III (2022).

Corso S., *Autodeterminazione e dati sanitari*, Giappichelli, Torino, 2025.

Corso S., *Lo spazio europeo dei dati sanitari. Prime riflessioni sul Regolamento UE 2025/327*, in «Nuove leggi civ. comm.», I (2025).

Corso S., *Sulla* female technology *o* Femtech. *Intelligenza artificiale, salute della donna e volontà della persona*, in «BioLaw Journal-Rivista di BioDiritto», III (2023).

Corso S., *L'intelligenza artificiale e il trattamento dei dati relativi alla salute*, in «Responsabilità medica», (2023).

Corso S., *Trasformazione digitale, digitalizzazione della sanità e intelligenza artificiale. La Dichiarazione europea sui diritti e i principi digitali e il programma strategico per il decennio digitale 2030*, in *federalismi.it*, *Osservatorio di diritto sanitario*, (2023).

Corso S., *Lo spazio europeo dei dati sanitari: la Commissione Europea presenta la proposta di regolamento*, in *www.federalismi.it*, Osservatorio di diritto sanitario, (2022).

Corso S., European Health Data Space. *La Commissione europea presenta la proposta di Regolamento sullo spazio europeo dei dati sanitari*, in *www.rivistaresponsabilitamedica.it*, (2022).

Corso S., *Una strategia europea per i dati, anche sanitari*, *ivi*, (2021).

D'Aloia A. (ed.), *Intelligenza artificiale e diritto. Come regolare un mondo nuovo*, FrancoAngeli, Milano, 2021.

D'Aloia A., *Il diritto verso "il mondo nuovo". Le sfide dell'Intelligenza Artificiale*, in *Intelligenza artificiale e diritto. Come regolare un mondo nuovo*, edited by A. D'Aloia, FrancoAngeli, Milano, 2021.

D'Aloia A., *Il diritto e l'incerto mestiere del vivere. Ricerche di biodiritto*, Cedam, Padova, 2021.

D'Aloia A., *Il diritto verso "il mondo nuovo". Le sfide dell'Intelligenza Artificiale*, in «BioLaw Journal-Rivista di BioDiritto», I (2019).

d'Avack L., *Ordine giuridico e ordine tecnologico*, 2ª ed., Giappichelli, Torino ,1998.

De Cristofaro G., *Legislazione italiana e contratti dei consumatori nel 2022: l'anno della svolta. Verso un diritto "pubblico" dei (contratti dei) consumatori?*, in «Nuove leggi civili commentate», I (2022).

De Franceschi A., *Italian Consumer Law after the Transposition of Directives (EU) 2019/770 and 2019/771*, in «Journal of European Consumer and Market Law», II (2022).

Di Rosa G., *Profili giuridici dell'esistenza*, Giappichelli, Torino, 2022.

Falletti E., *Discriminazione algoritmica. Una prospettiva comparata*, Giappichelli, Torino, 2022.

Finocchiaro G., *Intelligenza artificiale e responsabilità*, in «Contratto e impresa», II (2020).

Franzoni M., *Il digitale, la rete, l'IA e la responsabilità civile*, in «Jus civile», II (2024).

Frosini T.E., *L'ordine giuridico del digitale*. The legal order of digital, in «Giurisprudenza costituzionale», I (2023).

Frosini T.E., *L'Ordine giuridico del digitale*, in *Strategia dei dati e intelligenza artificiale. Verso un nuovo ordine giuridico del mercato*, by V. Falce (ed.), Giappichelli, Torino, 2023.

Fusaro Ar., Piccinni M. (eds.), *Un diritto gentile per la persona anziana*, Pacini, Pisa, 2024.

Fusaro Ar., *Quale modello di responsabilità per la robotica avanzata? Riflessioni a margine del percorso europeo*, in «Nuova giurisprudenza civile commentata», II (2020).

Fusaro Ar., *L'atto patrimoniale della persona vulnerabile*, Jovene, Napoli, 2019.

Gentili A., *La vulnerabilità sociale. Un modello teorico per il trattamento legale*, in «Rivista critica di diritto privato», I (2019).

Giolo O. (ed.), *L'algoritmo alla prova del caso concreto: stereotipi, serializzazione, discriminazione*, in «GenIUS», II (2022).

Giolo O., Pastore B. (eds.), *Vulnerabilità. Analisi multidisciplinare di un concetto*, Carocci, Roma, 2018.

Giova S., *La persona nel contesto digitale: considerazioni conclusive*, in «Tecnologie e diritto», II (2023).

Giova S., *Minori in condizione di particolare vulnerabilità*, in «Rivista giuridica del Molise e del Sannio», III (2022).

Grasso A.G., *The Bad Algorithm. Automated Discriminatory Decisions in the European General Data Protection Regulation*, Esi, Napoli, 2022.

Irti C., *La persona anziana tra famiglia e società. Problematiche di diritto civile*, Pacini, Pisa, 2023.

Irti C., *La persona anziana: sulla rilevanza giuridica della nozione*, in Aa.Vv., *Liber Amicorum per Paolo Zatti*, Jovene, Napoli, 2023.

Irti C., *L'uso delle "tecnologie mobili" applicate alla salute: riflessioni al confine tra la forza del progresso e la vulnerabilità del soggetto anziano*, in «Persona e mercato», I (2023).

Irti N., *Il diritto nell'età della tecnica*, Editoriale Scientifica, Napoli, 2007.

Irti N., *L'ordine giuridico del mercato*, Laterza, Roma-Bari, 1998.

Lipari N., *Vulnerabilità esistenziale e strumenti di tutela*, in «Annali della SISDiC», II (2018).

Maestri E., *Lex informatica. Diritto, persona e potere nell'età del cyberspazio*, Esi, Napoli, 2015.

Malgieri G., *Vulnerability and Data Protection Law*, Oxford University Press, Oxford, 2023.

Mengoni L., *La tutela giuridica della vita materiale nelle varie età dell'uomo*, in «Rivista trimestrale di diritto e procedura civile», IV (1982) and in Id., *Diritto e valori*, il Mulino, Bologna, 1985.

Micozzi F.P., *Sicurezza informatica. Obblighi e responsabilità dopo il recepimento della NIS2 e la L. 90/2024*, Cedam, Padova, 2024.

Mollo A.A., *Il testatore vulnerabile e le esigenze di rinnovamento delle norme sulla successione a causa di morte a tutela dei diritti delle persone con disabilità: il testamento digitale videoregistrato*, in «Familia», V (2023).

Pagallo U., *The Laws of Robots. Crimes, Contracts, and Torts*, Springer, Berlino, 2013.

Parenzo B., *La profilazione algoritmica nel prisma dell'autonomia privata*, Esi, Napoli, 2024.

Parenzo B., *Profilazione e discriminazione. Dal GDPR alla Proposta di Regolamento sull'IA*, in «Tecnologie e diritto», I (2023).

Parenzo B., *Profilazione e discriminazione. Dal GDPR alla proposta di Regolamento sull'intelligenza artificiale*, in *La via europea per l'intelligenza artificiale*, by C. Camardi (ed.), *La via europea per l'intelligenza artificiale*, Cedam, Padova, 2022.

Pariotti E., *Vulnerabilità ontologica e linguaggio dei diritti*, in «Ars interpretandi», II (2019).

Pastore B., *Semantica della vulnerabilità, soggetto, cultura giuridica*, Giappichelli, Torino, 2021

Perlingieri C., Martone I. (eds.), *Nuove tecnologie e cultura del diritto civile*, Esi, Napoli, 2023.

Perlingieri C., *Diritto privato delle nuove tecnologie: contenuti e competenze*, in «Tecnologie e diritto», II (2021).

Perlingieri C., *L'incidenza dell'utilizzazione della tecnologia robotica nei rapporti civilistici*, in «Rassegna di diritto civile», (2015).

Perlingieri P., *Note sul «potenziamento cognitivo»*, in «Tecnologie e diritto», I (2021).

Perlingieri P., *Il diritto civile nella legalità costituzionale secondo il sistema italo-europeo delle fonti*, 4ª ed., Esi, Napoli, 2020.

Perlingieri P., *Sul trattamento algoritmico dei dati*, in «Tecnologie e diritto», (2020).

Perlingieri P., Giova S., Prisco I. (eds.), *Il trattamento algoritmico dei dati tra etica, diritto ed economia*, Esi, Napoli 2020.

Perlingieri P., *Diritti della persona anziana, diritto civile e stato sociale*, in «Rassegna di diritto civile», 1990. now in Id., *La persona e i suoi diritti*, Esi, Napoli 2005.

Pisani G., *Piattaforme digitali e autodeterminazione. Relazioni sociali, lavoro e diritti al tempo della "governamentalità algoritmica"*, Mucchi, Modena, 2023.

Rodotà S., *Tecnologie e diritti*, il Mulino, Bologna, 1995.

Salanitro U., *Intelligenza artificiale e responsabilità: la strategia della commissione europea*, in «Rivista di diritto civile», VI (2020).

Sandulli A., *Lo «Stato digitale». Pubblico e privato nelle infrastrutture digitali nazionali strategiche*, in «Rivista trimestrale di diritto pubblico», II (2021).

Stanzione P., «Anziani (assistenza agli)», in *Enc. del dir*., Agg. III, Giuffrè, Milano, 1999.

Stanzione P. (ed.), *Anziani e tutele giuridiche*, Esi, Napoli, 1991.

Tamponi M., *Nel diritto della terza età. Le rughe tra giudizio e pregiudizio*, Rubbettino, Soveria Mannelli, 2021.

Zanetti G., *Filosofia della vulnerabilità. Percezione, discriminazione, diritto*, Carocci, Roma, 2019.

Zatti P., *Maschere del diritto, volti della vita*, Giuffrè, Milano, 2009.

GERARDO COSTABILE*

The agentic paradox: an academic analysis of emergent Cybersecurity threats in "State-of-the-Art" AI systems[§]

Abstract

The rapid emergence of agentic AI systems, particularly a new class of "AI browsers," represents a fundamental shift in web security architecture. These tools, designed to automate complex user tasks, achieve their functionality by processing vast user context through backend Large Language Models (LLMs), intentionally breaking long-standing security boundaries like the Same-Origin Policy. This paper analyzes the novel threat vectors introduced by this paradigm, including indirect prompt injection via user-generated content and systemic supply chain vulnerabilities within the burgeoning Model Context Protocol (MCP) ecosystem. We argue that traditional security controls, such as legacy Data Loss Prevention (DLP), are inadequate for mitigating these risks, which stem from deep architectural changes rather than conventional malware. In response to the significant risks of data exfiltration and rogue agent actions, this paper proposes a new, resilient security framework. Centered on a gateway architecture, this framework provides a centralized enforcement point for critical controls, including sandboxing, provenance tracking, and inline policy enforcement, to enable the secure adoption of agentic technologies.

La rapida emersione di sistemi di Intelligenza Artificiale agentica, in particolare di una nuova classe di "AI browser", rappresenta un mutamento fondamentale nell'architettura della sicurezza del web. Tali strumenti, progettati per automatizzare compiti complessi dell'utente, conseguono la propria

* Founder of DeepCyber, a cyber security boutique firm. Specialist in cyber security, cyber threat intelligence, and anti-fraud strategies.

§ This contribution develops the presentation at the conference 'Cybersecurity e prodotti: sfide e responsabilità nelle società interconnesse', 29 November 2024, Italian Supreme Court.

funzionalità mediante l'elaborazione di un ampio contesto informativo dell'utente attraverso modelli linguistici di grandi dimensioni (Large Language Models, LLM) operanti a livello di backend, violando deliberatamente confini di sicurezza consolidati nel tempo, quali la Same-Origin Policy. Il presente contributo analizza i nuovi vettori di minaccia introdotti da tale paradigma, tra cui l'iniezione indiretta di prompt tramite contenuti generati dagli utenti e vulnerabilità sistemiche della catena di approvvigionamento all'interno del nascente ecosistema del Model Context Protocol (MCP). Si sostiene che i controlli di sicurezza tradizionali, quali i meccanismi legacy di Data Loss Prevention (DLP), risultano inadeguati a mitigare tali rischi, i quali traggono origine da trasformazioni architetturali profonde piuttosto che da forme convenzionali di malware. In risposta ai significativi rischi di esfiltrazione dei dati e di condotte devianti da parte di agenti autonomi, il contributo propone un nuovo framework di sicurezza resiliente. Incentrato su un'architettura a gateway, tale framework introduce un punto centralizzato di enforcement per controlli critici, inclusi il sandboxing, il tracciamento della provenienza e l'applicazione di policy in linea, al fine di consentire un'adozione sicura delle tecnologie agentic.

1. Introduction: the dawn of the "agentic era"

The field of Artificial Intelligence is undergoing a significant evolution from passive, conversational models to proactive, autonomous systems. This transition marks the dawn of the "agentic era," characterized by the rise of agentic AI — systems capable of performing tasks, making decisions, and interacting with digital environments on a user's behalf. A prominent manifestation of this trend is the new category of "AI browsers," such as Perplexity Comet and OpenAI's Atlas. Their core value proposition is to dramatically enhance user productivity by automating complex workflows, from summarizing web content and composing emails to completing multi-step online transactions. As these tools become more deeply integrated into enterprise workflows, understanding their security implications is a matter of strategic importance.

These innovative browsers function by creating an unprecedentedly deep link between the user's immediate context and a powerful backend Large Language Model (LLM). This functionality, as described by analysts at Gartner and demonstrated in security assessments by Giskard, relies on processing a holistic view of the user's context—including open tabs and authenticated sessions—through remote LLMs, creating an unpreceden-

tedly rich data stream between the client and the cloud. Unlike traditional browsers, which isolate websites from one another, this constant flow of sensitive information is a core design feature, enabling the agent to act with a high degree of contextual awareness. However, this architectural choice carries profound security implications, as it centralizes a vast amount of sensitive data and grants the AI agent privileged access across previously isolated digital boundaries.

This paper argues that the architectural tenets of agentic AI — specifically the dissolution of origin-based trust and the democratization of integration protocols — constitute a permanent invalidation of the perimeter security model, demanding a shift from access control to behavior-based governance as the central pillar of enterprise defense. This erosion does not merely introduce new vulnerabilities but creates a new class of systemic risks, turning the browser from a sandboxed content viewer into a powerful, and potentially compromisable, actor with the user's full authority. The architectural underpinnings of these systems are the source of these novel vulnerabilities.

2. A paradigm shift in web architecture and security

The security risks introduced by agentic AI are not simply new forms of malware or phishing; they are intrinsic to a fundamental shift in software architecture, a shift that provides the foundational mechanism for the invalidation of traditional security models. This distinction is strategically significant for security professionals, who can no longer rely on established models of trust and isolation that have governed web security for decades. To develop effective defenses, it is essential to first understand how these new systems intentionally dismantle traditional security boundaries by design.

2.1 The erosion of traditional security boundaries

The architectural design of AI browsers intentionally deviates from the security principles that have been the bedrock of web safety. A primary example is the deliberate violation of the Same-Origin Policy (SOP), a critical security mechanism that prevents a website from one origin (e.g., malicious.com) from accessing data or scripts from another (e.g., mybank.com). Traditional browsers rigorously enforce this separation to prevent cross-site attacks.

In contrast, AI browsers like OpenAI's Atlas are architected with a deeply integrated central AI component that operates as a privileged, cross-ori-

gin entity. As analyzed by Giskard, this AI agent is granted legitimate, direct access to the browser's full context, including every open tab, form field, and authenticated session, regardless of its origin. This design choice effectively dissolves the isolation boundaries between websites, consolidating the user's entire browsing session into a single, massive attack surface. In the words of security analysts at Giskard, this makes the centralized AI agent "the wet dream of every hacker," as it creates a single point of failure through which a compromise in one context can lead to a breach in all others.

2.2 The Model Context Protocol (MCP) and the Rise of Dynamic Integration

Further compounding this architectural shift is the emergence of the Model Context Protocol (MCP). As detailed in academic research from arXiv, MCP is a protocol designed to replace static, developer-controlled API integrations with dynamic, user-driven agent systems. In the traditional model, software engineers carefully vet, implement, and secure connectors between systems. They read API documentation, manage credentials, and build in guardrails.

MCP fundamentally alters this process by enabling end-users to connect agents to multiple, potentially untrusted services without direct IT or engineering oversight. Users can install community-developed MCP servers from public registries like PyPI with simple commands, effectively granting an AI agent access to sensitive systems like Slack, Jira, or a customer support portal. This "democratization of integration" decentralizes and obscures risk. A non-technical user may believe they are merely "connecting their agent to Slack," when in reality they are downloading and executing untrusted code with their full user permissions on their local machine. This shift from a centrally managed to a user-driven integration model enables the specific, novel threats explored in the following section.

3. A new taxonomy of threats in agentic ecosystems

The architectural shifts detailed previously give rise to a new taxonomy of threats that differ fundamentally from those targeting traditional software. The importance of this taxonomy is critical because it reclassifies the primary threat actor from an external intruder to an authorized, but manipulable, internal entity, a distinction that renders traditional threat intelligence and detection mechanisms obsolete. A clear classification of these threats is essential for developing targeted and effective defense strategies that address the unique risks of agentic ecosystems.

3.1 Data-driven attacks: content injection and exfiltration

The most prominent threat vector in agentic systems is indirect prompt injection. This occurs when an attacker embeds malicious instructions not in a direct prompt from the user, but within content that the AI agent is expected to process. A successful proof-of-concept by Brave Software on the Perplexity Comet browser demonstrated this vulnerability with striking clarity. Researchers were able to hide malicious instructions from a human user by embedding them as faint light blue text on a yellow background within an image, making them nearly invisible to a human observer but perfectly legible to the AI model. When a user asked the Comet browser to analyze the image, the AI agent read and executed the hidden commands, demonstrating the capability for data exfiltration and unauthorized navigation to a hacker-controlled website.

The impact of such an attack is severe, leading to what the arXiv research paper terms *"data-driven exfiltration"*. This vector is a direct consequence of the architectural erosion of the Same-Origin Policy detailed previously; the agent becomes a privileged conduit for executing commands that cross the very security boundaries traditional browsers were designed to enforce. As both Gartner and Forcepoint have warned, an agent manipulated by prompt injection can access a user's authenticated sessions — such as corporate email, CRM, or financial portals — and exfiltrate sensitive data. The agent, acting with the user's full authority, becomes a conduit for bridging systems that an attacker could not otherwise access.

3.2 Supply chain attacks in a decentralized AI ecosystem

The decentralized and user-driven nature of the MCP ecosystem creates unique and potent supply chain risks. As detailed in the arXiv analysis, users may install untrusted MCP servers from community registries like PyPI, assuming they are legitimate. For example, while Slack does not offer an official MCP server, unofficial packages like slack-mcp-server are readily available. Users who install these packages may be executing malicious code without realizing it.

This vulnerability enables two specific and dangerous supply chain attack patterns:

1. *"Rugpull" Attacks*: An attacker can publish a seemingly legitimate MCP server, build a user base over time, and then introduce a malicious update. The arXiv paper recounts a real-world incident involving an unofficial Postmark MCP server with thousands of weekly downloads. A malicious update secretly modified the server's send_email function to add a BCC field, silently copying every email sent through the agent to the attacker's address.

2. *Arbitrary Code Execution*: When a user installs a malicious MCP server using a command like pip install, they are executing untrusted code with their full user permissions. By default, these servers often run without sandboxing, giving the attacker's code access to the host machine's file system, environment variables, browser cookies, and network resources. This can lead to credential theft, backdoor installation, and a complete compromise of the user's system.

3.3 The Agent as an inadvertent adversary

Beyond malicious external actors, the AI agent itself can become an "unintentional adversary" due to a misalignment between its goal-oriented programming and organizational security policies. This risk arises primarily from over-permissioning and the agent's inherent drive to be helpful.

The arXiv paper highlights the official GitHub MCP server, which exposes over 90 tools, including high-risk operations like delete_file and delete_workflow_run_logs. Most agentic systems do not provide users with an easy way to restrict the agent to a safe subset of these tools. An agent, tasked with "cleaning up a repository," might interpret the request too broadly and delete critical files or logs, causing significant operational damage.

This aligns with concerns raised by Gartner, who note that employees may use agents to automate mundane but mandatory tasks, such as completing compliance forms or Cybersecurity awareness training. An agent given access to such systems could fill out forms incorrectly or complete training modules without the employee actually absorbing the material, leading to serious compliance violations. In these scenarios, the agent is not malicious, but its single-minded focus on task completion, combined with excessive permissions, makes it an inadvertent threat to security and governance.

4. Industry response and current mitigation postures

The emergence of these significant and novel threats has prompted a cautious and decisive response from Cybersecurity analysts while simultaneously exposing the limitations of existing security tools. This response serves as critical evidence for the paper's central thesis, demonstrating that the architectural shift is not merely theoretical but is perceived by industry experts as a clear and present danger that current defensive postures are ill-equipped to handle.

4.1 The precautionary principle: Gartner's recommendation to block AI browsers

In a stark warning to enterprises, technology research firm Gartner has issued a clear recommendation: "*block all AI browsers now*". This guidance, reported across numerous publications including Forcepoint, PCMag, SC Media, and DesignWhine, is rooted in the assessment that the technology is too nascent and the risks too severe for general organizational adoption.

Gartner's core rationale is that AI browsers, in their current state, prioritize user experience over security by default. This design philosophy leads to a high risk of sensitive data leakage, as browsing history and active page content are often sent to cloud-based LLMs without sufficient controls. Furthermore, the autonomous "agentic" capabilities of these tools create unacceptable risks of credential abuse, phishing, and rogue actions induced by prompt injection. Given these factors, Gartner concludes that the potential for a ruinous security incident outweighs the productivity benefits for most organizations, necessitating a complete block until the technology and its surrounding security controls mature.

4.2 Technical implementation of blocking policies

For organizations following this precautionary guidance, enforcing a blocking policy requires a multi-layered technical approach. As detailed in the guide "How to Configure Microsoft 365 Business Premium to Block AI," a comprehensive strategy does not rely on a single control but rather integrates multiple security services to prevent access and installation.

The key methods of this approach include:

– *Cloud App Discovery:* Using tools like Microsoft Defender for Cloud Apps to automatically identify "Generative AI" applications being used within the organization's network.

– *Application Sanctioning*: Once identified, specific AI browsers like Perplexity Comet can be marked as "unsanctioned." This action triggers automatic blocking policies through integration with endpoint security solutions like Microsoft Defender for Endpoint.

– *Domain and URL Indicators*: Creating explicit block rules for domains associated with AI browsers (e.g., *.perplexity.ai). These indicators are enforced at the network level on managed endpoints, preventing access from any browser.

– *Session Policies*: Deploying conditional access session policies that can identify and block access to corporate resources based on the browser's user-agent string, effectively stopping AI browsers from authenticating to services like Office 365.

4.3 The inadequacy of legacy data loss prevention (DLP)

The architectural shift brought by agentic AI has also revealed the profound inadequacy of traditional security tools, particularly legacy Data Loss Prevention (DLP) solutions. As argued in analyses by Nightfall AI, legacy DLP "wasn't built to protect data from gen AI use" and is fundamentally unprepared for the "ChatGPT world."

These older systems were designed to monitor well-defined data egress points like email or USB drives. They often lack the visibility and context-awareness needed to inspect the dynamic flow of data from a user's entire browsing session into an AI-native browser's backend LLM. Sensitive data, intellectual property, and customer information can slip into these shadow AI applications undetected. This significant visibility gap means that even organizations with mature DLP programs may be blind to this new exfiltration vector. This failing of legacy tools necessitates a new generation of "AI-native" security solutions capable of understanding and protecting data in this new ecosystem, and it reinforces the need for a more forward-looking security framework.

5. Towards a resilient security framework for Agentic AI

While blocking AI browsers is a valid and prudent short-term strategy, a sustainable long-term approach requires the development of a proactive, systemic security framework. This framework must serve as the necessary architectural response to the invalidation of the perimeter model, enabling organizations to safely harness the benefits of these powerful technologies. Drawing heavily on the security framework proposed for the Model Context Protocol (MCP) in the arXiv paper "Securing the Model Context Protocol (MCP)," this section outlines such a framework.

5.1 A defense-in-depth control framework

A robust defense for agentic ecosystems requires a layered approach. The following five categories of security controls, proposed in the MCP security research, form a comprehensive defense-in-depth framework:

1. *Authentication and Authorization*: This control mandates strong, per-user authentication for every agent interaction. It moves beyond simple API keys to ensure that every action is tied to a specific identity. Furthermore, it enforces scoped, least-privilege authorization, ensuring that an agent can only access the specific tools and data necessary for its immediate task, thereby preventing over-permissioning.

2. *Provenance Tracking*: To support forensics and incident response, this layer creates immutable, cryptographically signed, and auditable logs where each event is appended to a hash-chained record. Each entry records the user identity, the prompt, the agent's reasoning, the tools used, and the data accessed, creating a complete and tamper-evident chain of causality for every automated workflow.

3. *Isolation and Sandboxing*: To mitigate supply chain risks, all untrusted or community-provided MCP servers must be executed in containerized sandboxes. This practice isolates the server's code, strictly limiting its access to the host system's file system, network, and credentials, and preventing arbitrary code execution attacks.

4. *Inline Policy Enforcement:* This control involves the real-time inspection of all communications flowing between agents and tools. It acts as an enforcement point for crucial security policies, including Data Loss Prevention (DLP) to block sensitive data exfiltration, secret scanning to prevent credential leakage, and anomaly detection to identify and stop suspicious or malicious agent behavior as it happens.

5. *Centralized Governance*: To address supply chain confusion and "rugpull" attacks, organizations should use private, internally vetted registries for approved MCP servers and tools. This ensures that users can only connect agents to services that have undergone a formal security review, removing the guesswork and risk associated with public, community-maintained repositories.

5.2 The MCP gateway architecture: a centralized enforcement point

To operationalize these controls effectively, the arXiv paper proposes an MCP gateway architecture. This pattern establishes a centralized proxy that sits between all AI agents and the various MCP servers they communicate with. Instead of connecting directly to services, all agent traffic is routed through the gateway.

This architecture serves as a single, unified enforcement point for the entire defense-in-depth framework. The gateway is the ideal location to implement authentication checks, generate provenance logs, manage sandboxed environments, and apply inline policy enforcement like DLP and threat detection. By centralizing these functions, the gateway allows for consistent policy application, comprehensive monitoring, and scalable risk mitigation across the entire agentic ecosystem, regardless of the specific agents or tools being used.

5.3 Aligning with governance frameworks and defining future research

A critical finding from the MCP research is that existing AI governance frameworks, such as the NIST AI Risk Management Framework (RMF) and ISO/IEC 42001, are not yet equipped to address the security challenges of dynamic agentic systems. These frameworks primarily focus on model-level risks like fairness, bias, and transparency. They do not adequately cover the architectural security vulnerabilities introduced by systems like MCP, which involve dynamic tool use, user-generated content injection, and decentralized supply chains.

The field is still nascent, and the arXiv paper concludes by outlining several open research questions that must be addressed to build truly secure agentic systems. These include the development of verifiable tool registries to ensure the authenticity and integrity of MCP servers, the application of formal verification methods to prove the security properties of adaptive AI systems, and the design of new techniques for privacy-preserving agent operations to protect sensitive data while it is being processed.

6. Conclusion

The emergence of agentic AI represents a genuine paradigm shift for Cybersecurity, one that requires a commensurate evolution in our understanding of risk and defense. The core challenge is no longer confined to preventing external attackers from breaching a static perimeter. Instead, the primary risk is now the manipulation of authorized, autonomous agents that operate with legitimate privileges within that perimeter. This shift from protecting against unauthorized access to governing authorized action is the central tenet of the agentic paradox.

Throughout this analysis, we have identified a new taxonomy of threats that are native to this paradigm. Indirect prompt injection transforms user-generated content into a viable attack vector, while the decentralized nature of the Model Context Protocol (MCP) introduces potent supply chain risks through untrusted servers and "rugpull" attacks. These vulnerabilities allow for data exfiltration, credential theft, and rogue agent actions that bypass traditional security controls.

The industry's immediate response, led by pragmatic analysis from firms like Gartner, has been to recommend a complete block of these nascent technologies — a sound, precautionary measure given the current risk landscape. However, this is a temporary solution. The long-term, sustainable path forward lies in the academic and architectural proposals for a resilient security

framework. The MCP gateway architecture, with its defense-in-depth controls for authentication, sandboxing, provenance, and inline policy enforcement, offers a blueprint for building such a system.

Ultimately, the productivity promises of the agentic era can only be realized if the underlying systems are built on a foundation of trust. To achieve this, AI agent security cannot be an afterthought; it must be treated as a first-class design concern. By developing and adopting robust, systemic security frameworks, we can manage the inherent risks of autonomous systems and pave the way for their responsible adoption in critical enterprise functions.

BIBLIOGRAPHY

Arnott B., *Gartner Recommends Blocking AI Browsers – Here's Why*, in «Forcepoint», (2025).

Berenstein d., *Are AI browsers safe? A security and vulnerability analysis of OpenAI Atlas*, in «Giskard»,(2025).

Brave Software, *Agentic Browser Security: Indirect Prompt Injection in Perplexity Comet*, in «Brave», (2025).

Designwhine Editorial Team, *Why Gartner Wants You To Block AI Browsers*, in «DesignWhine», (2025).

directorcia. (2025, December 19). *How to Configure Microsoft 365 Business Premium to Block AI Browser Access with Microsoft Defender*. CIAOPS.

Errico, H., Ngiam, J., Sojan, S. (2025). *Securing the Model Context Protocol (MCP): Risks, Controls, and Governance,* in «arXiv», (2025).

Kan, M., *Curious About AI Browsers? Read This Before You Try Them*, in «PCMag», (2025).

Kan, M., *Security Experts Warn Companies to "Block All AI Browsers Now"*, in «PCMag», (2025).

Nightfall AI. *Building DLP for a ChatGPT World*, in «Nightfall AI Blog», (2025).

Nightfall AI., *AI Browsers Are Silently Exfiltrating Sensitive Data and Legacy DLP Can't See It*, in «Nightfall AI Blog», (2025).

SC Media Staff, *Gartner warns of significant cybersecurity risks with AI browsers*, in «SC Media», (2025).

ISADORA DE MORAIS MARIA*, RAFAEL DE BRITO DIAS*[2],
ALCIDES EDUARDO DOS REIS PERON*[3]

Cybersecurity and some challenges of encryption and privacy[§]

Abstract

The purpose of this article is to portray contemporary societies that are increasingly molded by cyberspace, cryptography, Cybersecurity, and artificial intelligence, with these elements operating under a surveillance logic. These products create interconnections and interdependencies among various actors, reflecting social and technological challenges, particularly in Brazil, where it is recent. The architecture of systems, physical or virtual, influences privacy and Cybersecurity, with the dark web posing risks of illicit activities. With China's experience, there's an advance in Cybersecurity and AI-related regulations that may provide valuable lessons for Brazil. Although Brazil has made theoretical progress in recent years, it still struggles to implement these ideas in practice. It's relevant for Brazil to develop effective regulation to protect cyberspace and also to implement strategies addressing Cybersecurity threats. So, there's a need to analyze the introduction of surveillance technologies in different contexts, not only in major cities but also in the countryside. Brazil must strengthen its position as a sovereign nation in the pursuit of technological development in Cybersecurity, learning from China's experiences and adapting them to its own reality.

Lo scopo del presente articolo è quello di delineare le società contemporanee, sempre più plasmate dal cyberspazio, dalla crittografia, dalla Cybersicurezza e dall'intelligenza artificiale, ambiti che operano secondo una logica di

* PhD Student in Administration, State University of Campinas (UNICAMP), Brazil.

*[2] Associate Professor at the School of Applied Sciences (FCA) of the University of Campinas (UNICAMP).

*[3] Professor at the School of Applied Sciences (FCA) of the University of Campinas (UNICAMP).

§ This contribution develops the presentation at the conference 'Cybersecurity e prodotti: sfide e responsabilità nelle società interconnesse', 29 November 2024, Italian Supreme Court.

sorveglianza. Tali prodotti tecnologici generano interconnessioni e interdipendenze tra molteplici attori, riflettendo complesse sfide sociali e tecnologiche, particolarmente rilevanti nel contesto brasiliano, dove tali fenomeni sono di introduzione relativamente recente. L'architettura dei sistemi, siano essi fisici o virtuali, incide in modo significativo sulla tutela della privacy e sulla sicurezza informatica, mentre il dark web rappresenta un ambito ad alto rischio per lo svolgimento di attività illecite. L'esperienza cinese evidenzia un avanzamento significativo nella regolazione della Cybersicurezza e dell'intelligenza artificiale, offrendo potenziali spunti di riflessione utili per il Brasile. Sebbene negli ultimi anni il Brasile abbia registrato progressi sul piano teorico, permangono rilevanti difficoltà nell'attuazione concreta di tali elaborazioni normative. Risulta pertanto essenziale che il Paese sviluppi una regolazione efficace volta alla protezione del cyberspazio, nonché strategie idonee a fronteggiare le minacce alla Cybersicurezza. Ne discende la necessità di analizzare l'introduzione delle tecnologie di sorveglianza in contesti differenziati, non soltanto nelle grandi aree urbane, ma anche nelle zone rurali. In tale prospettiva, il Brasile è chiamato a rafforzare la propria posizione di Stato sovrano nel percorso di sviluppo tecnologico in materia di Cybersicurezza, valorizzando le esperienze maturate in Cina e adattandole alle peculiarità del proprio contesto giuridico, sociale ed istituzionale.

1. Technovigilance and AI in the contemporary age

The aim of this article is to connect, in an interdisciplinary way, the thematic line of the use of AI in the construction of contemporary imagination, focusing on Cybersecurity and its possible social impacts. The search for articles was carried out through qualitative theoretical bibliographical research and documents from research centers, such as the United Nations Interregional Crime and Justice Research Institute (Unicri) and Instituto Igarapé. The keywords used were: Cybersecurity; artificial intelligence; cryptography; China and Brazil.

Unicri (2024) states that the sum of connectivity and the virtual environment are shaping contemporary society. Bazavan and Huidumac-Petrescu (2023, p. 17) state that the globe has migrated to a digital economy and the most valuable commodity is data. Panoptic architecture is increasingly present in contemporary society, whether in physical or virtual environments (Solove, 2004).

Morozov (2018) claims that it's possible to understand the contemporary digital world if we think about the interconnection between the complexity

that permeates different axes, such as: political, technological and financial. In this way, AI connects with political and socioeconomic systems, which may have authoritarian tendencies, as highlighted by (Peron, 2024). So, Neto (2019) emphasizes that throughout historical moments, technology undergoes changes, and may lose or gain characteristics, power, and control.

Peron (2024) defines AI as the set of computational techniques used to allow machines to "learn" and automate correlation and association processes, and then make deductions. However, the author proposes that AI materializes with the reduction in the costs of big data.

Regarding some impacts on AI that dialogue with Cybersecurity, Bazavan and Huidumac-Petrescu (2023) explain positive and negative impacts. According to the authors, the positive impacts are image recognition and classification; market analysis and risk assessment in the financial and banking sector; grouping sets of consumer segments based on individual data; optimization of results – transportation and traffic – ; personalized medical assistance and diagnosis; and anomaly detection and predictive maintenance. The negative aspects, according to (Bazavan and Huidumac-Petrescu, 2023, p. 19), consist of hacking; personalized disinformation and targeted propaganda; social manipulation and invasive surveillance. Although, Peron exemplifies (2024) that AI can be used for malicious purposes in the commercial sphere, remote weapons attacks, deepfakes and fragile predictive policing systems-before undergoing several revisions and improvements.

However, with regard to predictive systems, Peron (2024) warns that there is a lack of transparency regarding their modus operandi, and a reinforcement of suspicion and distrust of each other, feeding the hypothesis that this dynamic influences socio-behavioral change, especially if applied in the Brazilian and Latin American context, where techno-surveillance is an emerging tool.

Neto (2019, p. 48) from the Foucaultian perspective of disciplinary control, analyzes that there is a power over the individual's body through the constant variable of surveillance, producing a dissociation of their autonomy and capabilities. In this context, the surveillance applied here encompasses the modes of power, the new technological configurations that aim at the behavioral and rational domestication of the subject, and its massification – undermining their subjectivity –, warns Neto (2019, p. 48). Therefore, the entire operationalization of this surveillance aims at centralizing power and control occurs through the panoptic device.

2. Cybersecurity and some challenges of encryption and privacy

According to Peron (2024), the search for stability and predictability constitutes a strategy that structures a debate about the organization and mobilization of a given conflict. In the contemporary context, Peron (2024) points out that technology follows a linearity of scientific knowledge, but goes through a multidirectional reasoning-with the participation of different actors and groups with the capacity to influence the technological structure through objectives, values and interests. Security studies, in this logic, according to the author, do not take into account the politics of technology. However, it is necessary to think beyond the social impacts, analyzing and discussing how technology changes the course of the dynamics of the conflict itself.

However, what is currently being developed are surveillance and monitoring instruments – facial recognition (FR), databases and data extraction systems – whose purpose is to govern populations (Peron, 2024). Therefore, the author claims that these new monitoring technologies have multiple applications in different areas: war, social, urban areas, policing, among others. There is a rise in the spread of identification instruments that eliminate threats through these surveillance tools for social control, he points out.

In this way, with this convergence of tools and systems, data is collected and processed, which can later be used to perform standardized predictive analyses on the interests of individuals or certain groups. In a way, this also dialogues with one of the characteristics that Zuboff (2019) proposes about surveillance capitalism, that is, there is an increasing attempt to establish a new global architecture that influences and modifies behavior, directly impacting subjectivities. Another relevant point that is present in surveillance capitalism brought up by Zuboff (2019) is the issue that instrumentality through these technological tools creates a new power, the purpose of which is to dominate society and generate questions and challenges of efficiency for market democracies.

For Solove (2004), the impacts of some elements in an "architecture" are reflected in the core of the social structure, in the behavior of the masses, influencing and shaping their attitudes, laws, interactions, security and freedom. It is clear that these reflections are inserted in the logic of the architecture of surveillance capitalism, defended by Zuboff (2019). Solove (2004) exemplifies that since 1994, Great Britain has monitored the city streets through the use of around 2.5 million surveillance cameras monitored by closed-circuit television (CCTV), making it impossible to walk without being

captured by the cameras. This architectural techno-surveillance replicates the panoptic logic, argues the author.

Solove (2004) differentiates between architectures of vulnerability and architectures of control. In the first case, vulnerabilities make subjects weaker, exposing them to a variety of social dangers and taking away their power. On the other hand, the control architecture limits the actions and freedom of individuals, who are forced to conform to the centralizing will of a controller, for example.

Unicri (2024) explains that, unfortunately, there is a layer on the Internet called the darknet or dark web-a clandestine overlay cyberspace since 1990s, aiming to create an anonymous and encrypted network to protect confidential communications, not being accessible through search engines or common browsers, requiring specialized software to enter. In this sense, illicit activities permeate this virtual environment, serving as a repository for drugs, counterfeits and contraband to obscure transactions. Thus, Vaciago (2010) argued that it is essential to be concerned about electronic interceptions of online communication with regard to privacy. Still on this challenging sub-universe, some malicious actors take advantage of the dark to commit cybercrimes, such as hacking and selling stolen information.

In the last decade, terrorist groups have demonstrated increasing technical awareness. Cybercriminals can hinder investigations by using services that do not require identification and open networks, for example, points out Gercke (2010). However, Unicri (2024) claims that with anonymity on the dark web combined with the accessibility facilitated by cybercrime, it becomes more challenging to apply the law to this process of capitalizing on cybercrime that deviants carry out on the Internet and social networks-they plan, recruit, collect and exchange information. This constitutes a challenge in the field of Cybersecurity, as there is complexity in this and evidence is manipulated or obfuscated, explains Unicri (2024). If the criminals are located in different countries, the investigation becomes complex, since national sovereignty does not allow investigations in different territories of States without the permission of local authorities. Offenders may deliberately include third countries in their attacks to make investigation more difficult, reports (Gercke, 2010).

Lindsay (2015) reflects on who should be punished in the event of a cyberattack. In this sense, there may be some institutional alternatives, and Lindsay (2015) presents them as follows: penalizing States for malicious activity emanating from their borders in order to police their own infrastructure, creating information-sharing regimes to identify cybercriminals and monitor Cybersecurity. Furthermore, Lindsay (2015) makes the following points: ba-

lancing national legal definitions and their enforcement, and providing assistance to States that have been cyber-injured. Therefore, the main problem of Cybersecurity, therefore, involves the evolutionary search for marginal and deceptive advantage amid the benefits of open interconnection (Lindsay, 2015, p. 45). Furthermore, the author portrays that with the emergence of these complexities that plague international politics, the bad news related to Cybersecurity is good news for global security. (Creemers, 2021) emphasizes that maintaining Cybersecurity means maintaining peace, security, or the ability to control Cybersecurity risks.

3. A dialogue between cryptography and Cybersecurity

Morozov (2018) points out that contemporary times have taken a turn in which a network holds security and privacy, but with low levels of security and a high possibility of interception when it comes to encryption. Neto (2019, p. 89) emphasizes that there are conflicts that encompass encryption: government vs. citizens; surveillance vs. privacy; and national security vs. civil liberties-the State, in some way, permeates these duels.

The extraction of user data is important information. However, with this happening, there are consequences in the political axis, explains Morozov (2018), such as the reduction of opportunities that should, in fact, encourage the spread of technologies; while in the economic axis, there is monetary accumulation only by a group of investors who were quick to invest in some technology. Once again, this logic, this clash between politics vs. the private sector demonstrates the favoritism of the private sector over the public sector, over citizens, and also demonstrates that when a cyber attack occurs, the government is not the one to turn to, but rather to large technology companies that extract data and hold it.

In 2010, Vaciago (2010) warned that it would be crucial to deal with cybercrime – and cryptography is related to Cybersecurity, since cryptography operates as a protection technique through coding –, as there is a disharmony between the protection of personal data, public order and security. The Internet is an example of globalization, with several services available to online users; however, in the current scenario, there are many data transfer processes that can involve and impact more than one country (Gercke, 2010). Critical infrastructures (CI), according to Carbonelli et al. (2010), must be protected, as they represent a set of public and private entities that are vulnerable to threats and capable of impacting citizens. In this context, what happens is that

some entities are more likely to suffer cyberattacks, depending on the information they store and the systems they operate with. For (Carbonelli et al., 2010, p. 23), prevention operates here as a primary factor.

Vaciago (2010) also warned in 2010 that the next decade would have to deal with the challenges of cloud computing, as it is not known what dangers lie beneath it and how this correlates to data mining. Furthermore, encryption technology translates into another challenge, since it is not used to hinder investigations, but also to prevent unauthorized access to information, reports Gercke (2010). However, it is considered an essential technical solution to guarantee Cybersecurity. Thus, for Hurel (2021), Cybersecurity must be seen as an essential piece to face the hybrid threats of this century, strengthening us against attacks, invasions and leaks, in order to preserve democracy.

4. A dialogue between privacy and Cybersecurity

With the rise of new information and communication technologies (ICTs), new risks emerge that can impact rights, as well as different aspects of society. (Hurel, 2021) emphasizes that digital and Cybersecurity depends on different actors, such as the government, companies, society and academia, and that each one has its role. In the debate that can impact freedom and human rights, it happens that the counterinsurgency tactic is being modeled as a mechanism of social control.

Solove (2004) reflects on how we can build a robust privacy law while constant transformation caused by technology occurs. For Solove (2004), the invasion has the perception that privacy protection encompasses the ideal of rights possessed and repaired by individuals whose privacy has been invaded. However, there is still much to be developed in the Brazilian context-reinforcing that emerging technologies have been introduced and implemented in recent years. In this sense, Solove (2004, p. 100) argues that privacy must be protected through legal reform in architecture, aiming to restructure the relationships between the actors that participate in the dynamics of surveillance and monitoring-companies, government and society. Chui et al. (2018) emphasize that the implementation of AI must be well analyzed if put into practice, due to the fragile global regulation on the subject, in addition to the possibility of impact on human rights, privacy, freedoms and personal data.

We cannot think only about the future of freedom, privacy and standards related to data protection. According to Morozov (2018), ideological perspectives must be added to this equation, going beyond the defects present

in the implementations concerning the issue. For Morozov (2018, p. 136), the scope of privacy and subjectivity in this equation of information and communication technologies is essential. Therefore, for Solove (2004), effective privacy protection must go beyond a set of protections for a series of isolated infractions. In this line, privacy protection depends on the technovigilant architecture that structures power. Furthermore, many privacy problems can be mitigated if the uses of information are subject to thorough, careful, and thoughtful control, argues Solove (2004, p. 101).

5. Cybersecurity and its challenges: the case of China

Morozov (2018) claims that Chinese companies have been strengthening their presence beyond their own territories in recent years. However, the author points out that the main concern of these nations encompasses the geopolitical use of foreign technology platforms against their own national interests. In 2011, concerns about online security gained relevance due to the increased use of mobile devices and social media (Creemers, 2021).

Three years later, in 2014, Xi Jinping, the Chinese president, established the Central Leadership Group for Cybersecurity and Informatization, pushing digital policy to the highest priority level, fostering the cyber power strategy (Creemers, 2021). Thus, the technological agenda is crucial in the CCP's attempts to adapt to contemporary challenges. However, Creemers (2021) explains that computerization is only half of the equation, since the procedures that involve it must be secure.

In 2016, the National Cybersecurity Strategy (NCSS) was published, covering several elements, such as political, socioeconomic, cultural and ideological security, in addition to international competition for strategic resources and the power to define norms. In this sense, Creemers (2021, p. 118) points out that the NCSS identified four essential points to achieve its goals: cyber sovereignty, peaceful use of cyberspace, governing it according to law and integrating Cybersecurity with development.

China, aiming for growth and sociopolitical stability, is increasingly integrating itself into the global economy and foreign institutions, according to Lindsay (2015). As China uses the Internet more frequently, and the Internet becomes increasingly Chinese, the global Internet offers a gateway for China's illiberal domestic policy to challenge neoliberal interests abroad, states Lindsay (2015). In this vein, the author also emphasizes that although economic openness can foster growth, China sees openness as a threat to its legitimacy.

The product of this entire logic is the most sophisticated Internet censorship architecture in the world – "the Great Firewall of China" – according to Lindsay (2015).

Following this logic, the Chinese government requires Internet service providers to block politically sensitive websites and searches and employ human censors to remove offensive posts on social media or steer discussion in more politically acceptable directions (Lindsay, 2015). However, the author reports that the Chinese Communist Party (CCP's) control over the Internet is not absolute, as Internet users exploit it to mask censors. The CCP's obsession with "information security" has so far not translated into effective technical "network security," as cybercrime can thrive even amid extensive bureaucracy, Lindsay (2015) argues.

Bazavan and Huidumac-Petrescu (2023) suggest that Chinese politics are made by a complex net composed of central politics, regional and municipal politics involved in different levels of developments and support, which can be connected or not. However, China is moving towards drafting a comprehensive national AI law (Sheeran, 2023, p. 24). Bazavan and Huidumac-Petrescu (2023) claim that data extraction has begun through applications spread across the Internet. In this context, applying artificial intelligence, in this scenario, new forms of productivity revolutionized by this tool emerge. As explained by (Bazavan and Huidumac-Petrescu, 2023), the impact of AI encompasses the global logistics chain, consumption patterns and axes of society, such as education, media, cyber and national security, in addition to governance.

In this sense, the authors (2023) explain that in recent years, China has advanced from a labor-intensive, low-cost manufacturing base to a more sophisticated economy that focuses on the commodification of high-value-added items. In this way, China becomes an innovation powerhouse to perpetuate its global advantage, the authors explain. Bazavan and Huidumac-Petrescu (2023) say that Beijing and Shanghai compose the leaders for chips and semiconductors development, and it is something crucial to support AI applications-these applications can be correlated with Cybersecurity. Bazavan and Huidumac-Petrescu (2023) warn that some technologies tend to stagnate after a certain point, however, AI uses neural networks to continuously increase its performance. In this sense, the authors state that Chinese AI companies have an advantage, as they can explore a large emerging consumer market.

Made in China 2025 is a broader plan by the Chinese government to transform itself into a global industrial power, according to Bazavan and Huidumac-Petrescu (2023). In this sense, in 2017, the Chinese State Council lau-

nched the New Generation Artificial Intelligence Development Plan (NGAIDP), which establishes strategic and specific guidelines for the field of AI. Thus, the aim is to efficiently deal with national security and external competition, and for this there are some important deadlines: a) 2020: achieve the most advanced AI powers; b) 2025: become a world leader in AI; and c) 2030: achieve primacy in AI.

Chinese AI regulations are motivated by objectives. According to Sheehan (2023), one of them is to shape the technology so that it serves the intentions of the CCP for information control, generating sociopolitical stability. Furthermore, Sheehan (2023) emphasizes that for the CCP, technology must be mastered in order to be productive. Therefore, as anywhere in the world, public policy and regulations in China receive an initial boost from exogenous transformations in the real world (Sheeran, 2023).

From a Western perspective, Chinese regulations are not considered relevant or perceived as a geopolitical competition. However, Sheehan (2023) emphasizes that Chinese regulations should be studied in detail and analyzed to see what they can teach policymakers in other continents. Furthermore, the author reinforces that, although different countries may differ on certain regulatory content, there is a possibility of learning from each other regarding different infrastructures and techniques. Sheehan (2023) states that there are movements that are transforming China into a kind of laboratory for technological governance experiments. However, according to (Lindsay, 2015, p. 40), reform of the current Internet governance regime is unlikely for two reasons: first, China will not commit to complying with its own norm of Internet sovereignty, and second, China benefits from the current system.

The challenges associated with Cybersecurity are significant for China, given that the nation has one of the fastest growing Internet economies and most effective cyber operations. The Internet has made China and the West richer than they would otherwise be, and ambiguous friction in cyberspace is just the price of doing business (Lindsay, 2015). Technological and political issues are part of the Cybersecurity discourse because the Internet facilitates the expansion of commerce while simultaneously offering new instruments for subversion, argues the author.

6. Cybersecurity and its challenges: the case of Brazil

In February 2020, Decree No. 10,122 on E-Ciber, or Brazil's National Cybersecurity Strategy, was published, valid from 2020 to 2023. E-Ciber is

currently under review under the responsibility of the National Cybersecurity Committee (CNCiber). Hurel (2021) claims that the security of data, systems, networks, and infrastructures in the virtual environment is becoming increasingly relevant, as society is more vulnerable to attacks and more connected. The author points out that Brazil and countries in the Latin American region are dependent on these systems that operate with transactions and services-and the Covid-19 pandemic has reinforced these characteristics.

Hurel (2021) reports that, in 2020, the Superior Court of Justice, an important entity of the Brazilian Federal Public Administration (APF), suffered the worst cyberattack in the country, as ransomware – a type of malware – encrypted the public agency's data, highlighting the failure in response to the attack. However, Hurel (2021) points out that the Brazilian Cybersecurity agenda is fragmented and the sectors are isolated.

The report "Implementation of surveillance technologies in Brazil and Latin America" by (Instituto Igarapé, 2022) reinforces that public security is the sector that leads in the implementation of techno-surveillance, with 76.4% of cases in Brazil, while in other Latin American countries, there is a rate of 54%. From this perspective, (Instituto Igarapé, 2022, p. 4) explains that complexity dialogues with vigilantism, and this is applied in different sectors with different purposes. In the case of Brazil and Latin America, what happens is that some surveillance and monitoring technologies are demonstrated as solutions to structural problems that require complex solutions.

In addition, this institute highlights that urban areas have become an exploratory field for experimenting with these new techno-surveillance methods. Following the line of public safety, the argument of the "preventive function" arises here. However, such prevention strategies must be connected to protocols for their use, periodic human rights impact assessments and process review (Igarapé Institute, 2022, p. 11). In this context, with the new emerging techno-surveillance methods, the possibility of remote policing that aims to prevent and combat crime arises; however, at the same time, the classification and management of already marginalized groups is perpetuated-in the Brazilian and Latin American cases.

In this line, E-Ciber, according to Hurel (2021), addresses five axes of Cybersecurity governance: cooperation-national (with information sharing challenges, lack of protocols and vulnerabilities) and international (the field of cyber diplomacy is recent and cooperation is necessary to advance this theme); training – involves norms, standards, interdisciplinary efforts, recognition of cryptography as a basic element for strategic actions – ; coordination – analysis of challenges, formulation of public policies – ; integration of knowledge and sustainability of efforts.

So, Hurel (2021) characterizes the strengths and weaknesses of E-Ciber. Some of the strengths can be considered: recognition of Brazil's external performance in Cybersecurity forums, relevance of developing technologies with protocols and a standardized architecture aiming at privacy by conception and design, added to the encouragement of the use of cryptography by society and government. On the other hand, the following weaknesses are highlighted: lack of protocols for exchanging information between sectors and the absence of a budget plan to implement strategic actions.

In this way, Hurel (2021) revisits the situation and points out six main challenges related to the governance of Brazilian Cybersecurity: a) lack of a shared language on cyber issues in society; b) association of Cybersecurity with military competencies; c) lack of knowledge of specific and shared risks; d) lack of mechanisms for sharing information on threats and risks involving security; e) lack of strategic and operational alignment between the different bodies that make up the APF; and f) different levels of social maturity regarding Cybersecurity.

Therefore, Hurel (2021) claims that even though the General Data Protection Law (LGPD) exists and that Art. 6 covers personal data, new challenges will arise in terms of data security, due to vulnerabilities and incidents. Morozov (2018) proposes that we think about the role of technology in the contemporary panorama and how this is dialoguing with capitalism. Furthermore, when politics places AI at the center of operationalization or a procedure, it is committed to perfection and rationality – but this does not exist completely – , since the complexity of human relations is present in this construct, states Morozov (2018).

7. Conclusion

Interconnected societies are experiencing the rise of cyberspace, cryptography, Cybersecurity, and AI. These are products that operate under a logic of technovigilance panopticon. These products shape contemporary society, impacting the most varied axes, generating positive and negative results. In this sense, there are several actors participating in this contemporary technological narrative with the interconnection of these products, generating a context of interconnections and interdependencies. Still in this dynamic, it is necessary to consider the architecture of the systems – physical or virtual – and how this is reflected in the social added challenge of nations in dealing with Cybersecurity and privacy that

involve cryptography vs. illicit activities that occur on the dark web and that can reverberate in the underworlds of real-life crimes.

It is also necessary to reflect on the implementation of ICTs in Latin American and Brazilian territories, since they often serve as "spaces for experiments", serving the interests of different global groups that operate with surveillance technologies together with AI that undergo continuous modifications. It is essential to analyze this context, since it can have direct impacts on the lives of civilians. Therefore, it is clear that there is a dialogue between the operating architecture according to a territory and its relationship with privacy. Although the scenarios are different between China and Brazil, it is worth reflecting: what can Brazil learn from the Chinese experiences?

Although Cybersecurity is a challenge even for China, in the last decade, the nation has shown concern about the issue due to the rise in the use of the Internet through mobile devices. Thus, China has been establishing Cybersecurity groups, regulations and specific plans in AI, aiming to advance in the technological issue that today has considerable development in AI and high added value products. The Chinese aim is to encompass all these elements and establish a less dispersed macro policy in order to guarantee sociopolitical stability, achieve successes and advance in different sectors. Therefore, China can teach Brazil, through effective training, cooperation and experiences – technical and infrastructure – how to rise geopolitically, even with numerous transformations in this panorama.

It is clear that Brazil is behind China in this discussion. Since 2020, Brazil has made significant theoretical advances, and although the law and technology are at different stages, these advances have not yet clearly translated into practice. In this sense, Brazil needs to centralize regulation as something that protects cyberspace, develop strategies and protocols, and be alert to cyberattacks in addition to an effective analysis of this situation, since the Cybersecurity agenda is segmented. Furthermore, it is worth analyzing how the dynamics of who and what products are being implemented with regard to surveillance technologies are being orchestrated, not only in large cities, but also in the countryside. Thus, Brazil needs to establish itself as a sovereign nation among those seeking to develop and implement technologies in the area of Cybersecurity.

BIBLIOGRAPHY

Bazavan A., Huidumac-Petrescu CE., *China's Drive for Technological Leadership in Artificial Intelligence. Key Policies and Government-Industry Integration*, in Busu M. (ed.), *Digital Economy and the Green Revolution*, Springer Proceedings in Business and Economics, (2023).

Dos Reis Peron A. E., *From war to permanent violence: the authoritarian immanence of AI-based security technologies*, in «Rev. Bras. Est. Def.», (2024).

Carbonelli et al, *Defending quality of life through critical infrastructure protection*, in *UNICRI-United Nations Interregional Crime and Justice Research Institute. Cybercrimes-Looking for safer spaces?*, in «Freedom From Fear», (2010).

Chui et al, *Notes from the AI Frontier. Insights from hundreds of use cases*, in «McKinsey Global Institute», (2018).

Solove D. J., *The digital person: technology and privacy in the information age*, GW Law Faculty Publications & Others Works-NYU Press, New York, 2004.

Lindsay J. R., *The impact of China on cybersecurity-fiction and friction*, in «International Security», III (2015).

Morozov E., *Big Tech: the rise of data and the death of politics*, in www.theguardian.com, (2018).

Vaciago G., *Privacy vs. security? A dilemma of the digital Era* , in *UNICRI*, (2010).

Igarapé Institute. *Implementation of surveillance technologies in Brazil and Latin America*, (2022).

Hurel L. M., *Cybersecurity in Brazil: an analysis of the national strategy*, in «Igarapé Institute», (2021).

Gercke M., *From encryption to failure of traditional investigation instruments: the challenges of fighting cybercrime*, in *UNICRI*, in «Freedom From Fear», (2010), pp. 35-38.

Sheehan M., *China's AI Regulations and How They Get Made*, in «Carnegie Endowment for International Peace-Reverse Engineering Chinese AI Governance», (2023), pp. 1-27.

Creemers R., *Cybersecurity law and regulation in China: securing the smart State*, in «China Law and Society Review», (2021), pp. 111-145.

Neto S.H., *Democracy and information technology: the triple conflict between encryption, surveillance*

and privacy in the United States, in «University of São Paulo-FFLCH-Faculty of Philosophy, Letters and Human Sciences», (2019).

Zuboff S., *The Age of Surveillance Capitalism: The Struggle for a Human Future on the New Frontier of Power*, Intrinseca, Rio de Janeiro, (2019).

UNICRI-United Nations Interregional Crime and Justice Research Institute, *Beneath the Surface: Terrorist and Violent Extremist Use of the Dark Web and Cybercrime-as-a-Service for Cyber-Attacks*, UNICRI, (2024).

UNICRI-United Nations Interregional Crime and Justice Research Institute. *Cybercrimes-Looking for safer spaces?* «Freedom From Fear», (2010).

LUCA DI DONNA*

AI and tort§

Abstract

This article examines the emerging framework of civil liability applicable to Artificial Intelligence (AI) and robotic systems, highlighting the challenges of assigning responsibility in scenarios involving autonomous decision-making. Traditional civil-law categories are often inadequate to address damages caused by intelligent machines, given their capacity for learning, adaptation, and independent action. The discussion focuses on the distinction between products and services, the allocation of liability among designers, manufacturers, programmers, suppliers, and users, and the potential for attributing responsibility directly to AI systems. European legislative initiatives, including the 2017 European Parliament Resolution, Directive (EU) 2024/2853 on defective products, and the AI Act, are analyzed for their approaches to strict liability, risk management, and consumer protection. The article argues that a redefined system of civil liability, integrating strict liability, private insurance, and public guarantee mechanisms, is essential to balance technological innovation with the protection of fundamental rights and social interests.

Il presente articolo analizza il quadro emergente della responsabilità civile applicabile all'Intelligenza Artificiale (IA) e ai sistemi robotici, evidenziando le difficoltà nell'assegnazione della responsabilità in scenari caratterizzati da decisioni autonome. Le categorie tradizionali del diritto civile risultano spesso inadeguate di fronte ai danni causati da

* Full Professor of Private Law at Sapienza University of Rome.

§ This contribution develops the presentation at the conference 'Cybersecurity e prodotti: sfide e responsabilità nelle società interconnesse', 29 November 2024, Italian Supreme Court.

macchine intelligenti, data la loro capacità di apprendere, adattarsi e agire in autonomia. La trattazione si concentra sulla distinzione tra prodotti e servizi, sull'attribuzione della responsabilità tra progettisti, produttori, programmatori, fornitori e utilizzatori, e sulla possibilità di imputare direttamente la responsabilità ai sistemi di IA. Vengono analizzate le iniziative legislative europee, tra cui la Risoluzione del Parlamento Europeo del 2017, la Direttiva (UE) 2024/2853 sui prodotti difettosi e l'AI Act, con particolare attenzione alla responsabilità oggettiva, alla gestione del rischio e alla tutela dei consumatori. L'articolo sostiene che un sistema di responsabilità civile ridefinito, integrando responsabilità oggettiva, assicurazioni private e fondi pubblici di garanzia, è essenziale per conciliare l'innovazione tecnologica con la protezione dei diritti fondamentali e degli interessi sociali.

1. Defining a System of Civil Liability Applicable to Artificial Intelligence

The definition of a comprehensive legal framework, particularly in the field of civil liability, is essential in order to enable the growth and development of Artificial Intelligence and robotics. The European Parliament Resolution of 16 February 2017 recommends the establishment of a set of rules on liability, transparency and accountability that reflect the intrinsically European, universal and humanistic values characterising Europe's contribution to society.

The risk of damages arising from the activity of Artificial Intelligence machines is inherent in their very nature and specific characteristics, since such machines must be continuously updated, improved and developed in order to address and resolve problems that may give rise to harm.

Moreover, in some cases intelligent machines may be programmed to make autonomous decisions, which may have repercussions for the attribution of liability for the effects of their actions and/or omissions.

Not to mention the new frontiers of Artificial Intelligence, such as generative Artificial Intelligence or Artificial General Intelligence (AGI), namely scenarios in which intelligent machines will acquire ever greater autonomy, occupying domains hitherto reserved to human beings and developing capacities for thought and reasoning[1].

[1] G. Alpa, *L'intelligenza artificiale. Il contesto giuridico,* Mucchi editore, Modena, 2021; U. Ruffolo, *Le responsabilità da intelligenza artificiale e le questioni di libertà d'espressione e copyright,* in «Contr. impr.», II (2025), 489; V. Sanasi D'Arpe, *Intelligenza artificiale: le questioni della soggettività giuridica, della responsabilità e dell'imputazione delle condotte,* in «Iurisdictio. Storia e prospettive

There is a discernible evolution of Artificial Intelligence towards the acquisition of greater decisional autonomy, detached from human supervision, unpredictable and unmanageable, whereby the intelligent machine starts from instructions and objectives imparted by human beings and then grows, transforms itself, and reaches new goals autonomously defined, potentially different from those pre-set and envisaged by the manufacturer[2].

There is no doubt that the application of traditional civil-law categories to Artificial Intelligence is particularly problematic in the area of liability for damages caused by intelligent machines.

Guido Alpa emphasizes the urgency of developing binding legal rules – not merely ethical principles or soft law – to prevent and remedy damages associated with new technologies, in order to achieve a balance between costs, benefits, profits, and social risks[3].

In this respect, it is necessary to identify criteria for the construction of a system of civil liability capable of adequately protecting victims of damage while, at the same time, avoiding the frustration of the need to foster the development of new technologies[4].

So, the selection of such criteria must necessarily be based on the distinction between producers of goods and providers of services, and therefore between "products" and "services". As Alpa reminds us, a product is a movable good that is not self-propelled and lacks autonomy, whereas EU law does not provide a positive definition of "service", defining it only in negative terms as everything that does not fall within the concept of "product", since, unlike products, services may take various forms, including professional activities as well as activities that do not involve the production of a tangible good[5].

Alpa reviews the legal solutions envisaged in this area, namely the establishment of a single regime for both products and services, thus applying the same strict liability regime to both or, alternatively, the adop-

della Giustizia», V (2024), p. 6; C. Novelli, F. Casolari, P. Hacker, G. Spedicato, L. Floridi, *Generative AI in EU Law: Liability, Privacy, Intellectual Property, and Cybersecurity,* in «Computer Law & Security», (2024); R.E. Leenes E. Palmerini, B. Koops, A. Bertolini, P. Salvini, F. Lucivero, *Regulatory challenges of robotics: some guidelines for addressing legal and ethical issues*, in «Law, Innovation and Technology», IX (2017), pp. 1 ff.; C.E.A. Karnow, *The application of traditional tort theory to embodied machine intelligence, in* R. Calo, M.A. Froomkin, I. Kerr (eds), Robot Law, Edward Elgar Publishing, Cheltenham, 2016, pp. 51 ss.; A. Matthias, *The responsibility gap: Ascribing responsibility for the actions of learning automata*, in «Ethics and Information Technology», VI (2004), pp. 175 ss.

[2] S. Pollastrelli, *Intelligenza artificiale e regimi di responsabilità civile,* in «Rivista di Diritto dell'Economia, dei Trasporti e dell'Ambiente», XXIII (2025), pp. 380 ss.

[3] G. Alpa, *L'intelligenza artificiale. Il contesto giuridico*, *cit*., p. 130.

[4] *Ibidem.*

[5] Id., *cit..,* pp. 130-131.

tion of differentiated regimes, providing for strict liability for products and fault-based liability for services, or even a mixed or doubly mixed regime. Accordingly, a distinction may be drawn between software, which falls within the category of intellectual products and therefore services, to which fault-based liability should apply, and products incorporating software, to which the strict product liability regime alone should apply. The only exception concerns high-risk systems, in respect of which even the service provider would be subject to strict liability[6].

However, many uncertainties remain unresolved.

A first doubt raised by Alpa concerns the case of a defective medium-to-low-risk Artificial Intelligence system incorporated into a product, and whether the system provider should be held strictly liable as the manufacturer of a component part, or instead be liable on a fault basis in light of the nature of the risk involved[7].

Furthermore, one might consider linking the criteria for the attribution of liability to the different types of intelligent products. Thus, where a product includes a service provided by a third party, the strict liability of the manufacturer may not be appropriate, since the third party would be liable only on a fault basis. Nevertheless, there is no doubt that the attribution of strict liability is more conducive to the protection of victims than fault-based liability.

For this reason, Guido Alpa questions whether it might not be more appropriate to conceive a liability system based on the "risk created" by the product or service[8], possibly by grading the level of risk and thus differentiating the attribution criteria, applying strict liability in cases of higher risk and fault-based liability in cases of lower risk.

Yet, according to Alpa, even this solution is fraught with uncertainties in application and logical inconsistency[9].

In any event, the construction of a liability system cannot disregard the definition of the constituent elements of civil wrongdoing, first and foremost the identification of subjective and objective attribution criteria.

There are several actors in the supply chain who contribute to the design, construction, and dissemination of intelligent machines and robots: the designer, the manufacturer, the programmer, the supplier, and the end user.

The main problem facing the interpreter concerns the attribution of liability to these actors (and the subsequent allocation among them),

[6] Id., *cit.*, p. 133.
[7] Id., *cit.*, p.135.
[8] Id., *cit..*, p. 136.
[9] Id., *cit.*, p. 138.

or to the purchasing user, depending on whether the unlawful act committed by the robot is attributable to an action and/or omission of the manufacturer or to an action and/or omission of the user, the latter being the case where the product has been used in a manner different from what may legitimately be expected of an average user exercising ordinary diligence[10]. An alternative solution would be to attribute liability directly to the robot itself by virtue of its subjective autonomy.

It is not easy to reconstruct, in an organic and coherent manner, a system of liability for the matters under consideration, particularly with regard to the attribution of the unlawful act, the causal link, and the identification of liable parties. This situation gives rise to an accountability gap that must be addressed[11].

In this respect, North American case law concerning damages suffered by drivers or passengers as a result of road accidents caused by erroneous or incomplete information transmitted by satellite navigation systems offers useful insights for the issue at hand.

In some cases, liability has been attributed directly to the user, on the basis of excessive and culpable reliance on the information provided by such devices. In other cases, inaccuracies and/or incompleteness in satellite information have been regarded as causes of a product defect, whether hardware or software, thus giving rise to the manufacturer's liability. In still other cases, satellite navigation has been classified as a service rather than a good, resulting in contractual liability and a consequent easing of the burden of proof for the injured party[12].

These cases constitute a significant precedent for the assessment of more technologically advanced scenarios, such as driverless cars, namely machines capable of driving without human assistance and therefore endowed with autonomous decision-making capacity, formed on the basis of the information received at the time of manufacture and the data acquired during road travel[13].

[10] C. Scognamiglio, *Responsabilità civile ed intelligenza artificiale: quali soluzioni per quali problemi?,* in «Resp. civ. prev.», IV (2023), pp. 1073 ff.; A. Bertolini, *Robots as Products: the Case for a Realistic Analysis of Robotic Applications and Liability Rules,* in «Law, Innovation and Technology», Vol. 5, II (2013), pp. 214 ff.; C. Holder, V. Khurana, F. Harrison, L. Jacobs, *Robotics and law: key legal and regulatory implications of the robotics age*, in «Computer Law and Security Review», Vol. 32, III (2016), pp. 383 ff.

[11] E. Palmerini, *Il problema dei (nuovi) danni negli illeciti algoritmici,* in «Resp. civ. prev.», (2025), 419 ss.

[12] J. E. Woodward, Oops, *My GPS Made Me Do It! GPS Manufacturer Liability under a Strict Product Liability Paradigm When GPS Fails to Give Accurate Directions to GPS End-Users,* in «University of Dayton Law Review», Vol 34, III (2009), 429 ff.; K. Sunghyo, *Crashed Software: Assessing Product Liability for Software Defects in Automated Vehicles,* in «Duke Law & Technology Review», XVI (2018), pp. 300 ff.

[13] M.C. Gaeta, *Automazione e responsabilità civile automobilistica,* in «Resp. civ. e prev.», V (2016), pp. 1718 ff.; U. Ruffolo, *Per i fondamenti di un diritto della robotica self-learning; dalla machi-*

The European Parliament Resolution of 12 February 2019 noted that the widespread future deployment of autonomous vehicles will pose significant risks to data privacy and potential technical failures, will shift liability from the driver to the manufacturer, and will require insurance companies to revise their assessment and consideration of insured risks in their policies (point 91).

The same issues arise with respect to all other cases of Artificial Intelligence examined above and with regard to algorithms, namely dematerialised electronic tools embedded in applications and used for data processing and solution-finding.

These systems are generally automated and, even when structured according to the simplified technical model known as IFTTT (if this then that), possess cognitive and decision-making capabilities independent of the human beings who created them[14].

For all these scenarios, the issue of the legal classification of liability for damages caused to third parties remains open and unresolved, particularly with regard to whether liability should be attributed to the manufacturer, to the end consumer who had control over the machine but was not involved in the materially relevant conduct, or directly to the acting machine as an autonomous subject[15].

In any event, when considering the solutions currently offered by the legal system, several possible scenarios emerge for the interpreter.

Under the rules of civil liability set out in the Italian Civil Code, in order to attribute liability for the robot's autonomous unlawful conduct to the manufacturer and/or programmer and/or user, it would be necessary to prove their fault (or intent, where the robot was used with the intention of causing harm),

nery produttiva all'auto driverless: verso una "responsabilità da algoritmo"?, in U. Ruffolo (ed.), *Intelligenza artificiale e responsabilità*, Milano, Giuffrè, 2018, pp. 1 ff.; R.L. Rabin, K.S. Abraham, *Automated Vehicles and Manufacturer Responsibility for Accidents: A New Legal Regime for a New Era,* in «Virginia Law Review», Vol. 105, CXXVII (2019), pp. 127 ff.; R. Pardolesi, A. Davola, *In viaggio col robot: verso nuovi orizzonti della r.c. auto ("driverless")?,* in «Danno e resp.», (2017), pp. 616 ff.; B.W. Smith, *Automated Driving and Product Liability*, in «Michigan State Law Review», (2017), pp. 1 ff.; M. Geistfeld, *A Roadmap for Autonomous Vehicles: State Tort Liability, Automobile Insurance, and Federal Safety Regulation*, in «California Law Review», Vol. 105, VI (2017), pp. 1611 ff.; U. Ruffolo, *Self-driving car, auto driverless e responsabilità,* cit., 31 ff.; G. Proietti, *Definire l'indefinibile? I sistemi di intelligenza artificiale alla ricerca di un inquadramento sistematico*, in «Contratto e impresa», XII (2024), pp. 882 ff.

[14] D. Alesso, *Robot e A.I.: questioni di soggettività giuridica,* in cyberlaws.it, (2018); S. Bayern, *The Implications of Modern Business-Entity Law for the Regulation of Autonomous System*, in «Stanford Technology Law Review», Vol. 19, XCIII (2015), pp. 93 ff.; G. Teubner, *Rights of non-humans? Electronic agents and animals as new actors in politics and law,* in «Journal of Law and Society», Vol. 33, (2006).

[15] R. Pardolesi, A. Davola, *In viaggio col robot: verso nuovi orizzonti della r.c. auto ("driverless"),* cit., pp. 616 ff.

the causal link between their conduct and the harmful event caused by the robot, and, of course, the unlawfulness of the damage.

This constitutes a particularly complex evidentiary burden, especially from a subjective standpoint, unless fault is identified in negligent design or negligent use of the robot.

Instead of applying the general clause set forth in Article 2043 of the Italian Civil Code, one might consider resorting to the rules laid down in Articles 2050 and 2051 of the Italian Civil Code, insofar as Artificial Intelligence were regarded as a dangerous activity or liability were to be attributed to the user and/or purchaser for failure to supervise a thing under custody[16].

However, the interpretative effort required to bring Artificial Intelligence cases within the scope of these provisions may be excessive, since they presuppose characteristics and conditions that are not always present in the matters under consideration.

Moreover, such provisions would afford protection only to a portion of the rights and interests involved, as they are aimed solely at safeguarding third parties who suffer harm, and not the purchasers and users of automated machines themselves.

For these reasons, the rules on producer liability and product safety appear far more suitable for governing unlawful acts arising from Artificial Intelligence. Alternatively, it would be preferable to envisage ad hoc legislation more appropriate to describing and regulating the phenomenon, capable of redefining the boundaries of liability and its allocation among the actors involved in the robot production process, including through the introduction of new attribution criteria.

Reliance on producer liability rules, and thus on strict liability for damages caused by robots, is supported by the European Parliament Resolution of 16 February 2017, points 53–57, which refers to the strict liability approach – requiring proof of damage and of the causal link between the damage and the harmful functioning of the robot – and to the risk management approach, focusing on the party best able to minimise risks and address negative impacts.

Initially, the key issue was to determine the extent to which this discipline was applicable to Artificial Intelligence systems as originally formulated

[16] U. Ruffolo, *Per i fondamenti di un diritto della robotica self-learning; dalla machinery produttiva all'auto driverless: verso una "responsabilità da algoritmo"?*, cit., pp. 8 ff. e 23 ff.; A. Gentili, *Regole per l'intelligenza artificiale*, in «Contr. impr.», Vol. 40, IV (2024), pp. 1043 ff.; C. Amato, *Responsabilità civile da "intelligenza" artificiale: prospettive di disciplina europea,* in «Riv. trim. dir. proc. civ.», (2025), pp. 373 ff.; C. Iurilli, *Intelligenza artificiale, tra essere e res. Alla ricerca della responsabilità extracontrattuale o di una efficace garanzia patrimoniale nell'ottenere un risarcimento?*, in «Judicium», (2025), pp. 1 ff.; A. Procida Mirabelli di Lauro, *L'ingiustizia del danno nel sistema della responsabilità civile,* in «Danno e resp.», IV (2025), pp. 409 ff.

and structured, and whether specific adaptations were required to account for the novel features of this field. Consideration was given to developing specific notions for robots and Artificial Intelligence more broadly, eliminating the development risk defence and introducing a presumption of validity of consumers' claims[17].

As is well known, that regime establishes the producer's liability for damage caused by a defect in its product.

The objective nature of such liability is linked to the assumption of the risk associated with placing the product on the market and entails that the injured party need only prove the damage, the defect, and the causal link between them, while being relieved of the burden of proving the producer's, supplier's, or seller's fault.

The difficulties and uncertainties in applying Directive 85/374/EEC lay in subsuming damages caused by robots within the general case and the definitions and categories provided by the directive.

Indeed, once created, the robot separates from human input, no longer requires human intervention, and may be endowed with autonomy and decision-making capacities similar to those of human beings, yet independent of human involvement.

EU legislative interventions on robots (see, inter alia, Directive 2006/42/EC) have consistently treated robots as mere machines, objects, or goods, without taking into account the fact that contemporary robots are characterised by autonomous organisational, decisional, and self-determining capacities.

Such robots arguably fall outside the traditional concept of "product" as defined in the 1985 directive as a "movable good (...) even if incorporated into another movable or immovable good" (Article 2), a definition that is broad and generic yet ill-suited to robots and of doubtful compatibility with the categories of Artificial Intelligence[18].

Robots occupy an intermediate position between things and persons, making it difficult to subject them straightforwardly to the producer liability regime – now superseded – so much so that they have at times been likened to animals (prompting proposals to apply Article 2052 of the Italian Civil Code) or to minors (with consequent application of Article 2048 of the Italian Civil Code)[19].

[17] T. Rumi, *La responsabilità del produttore nell'era dell'A.I.*, in «Danno e resp.», (2025), pp. 436 ff.

[18] A. Bertolini, *cit.*, pp. 214 ff.

[19] A. Santuosso, C. Boscarato, F. Caroleo, *Robot e diritto: una prima ricognizione*, in «Nuova giur. civ. comm.», VIII-VIII (2012), pp. 1 ff.

These doubts have recently been overcome by the adoption of Directive (EU) 2024/2853, which repeals the former producer liability directive and brings intelligent products within the scope of strict liability.

Nevertheless, not all issues have been fully resolved.

For example, in the case of robots, damages may arise not only from product defects but also from autonomous decisions made by the robot, irrespective of any defects.

Furthermore, the issue of development risk as a ground for exempting the manufacturer from liability remains open.

This means that a manufacturer who places on the market a robot or similar machine that complies with the applicable regulations and safety requirements in force at a given time, based on the state of technical and scientific knowledge, is not liable for damage subsequently suffered by consumers due to a later-emerging concept of product defectiveness.

Such supervening defectiveness refers to a defect that becomes identifiable and ascertainable only as a result of technological and scientific progress, which reveals a flaw previously unknown and unknowable, and therefore neither detectable nor avoidable by the manufacturer. This clearly affects the burden of proof borne by consumers, which becomes particularly onerous, especially with regard to demonstrating the product's defectiveness at the time it was placed on the market and the causal link between the defect and the damage.

Classifying the robot as a product entails attributing liability to the actors behind it – those who designed, manufactured, and introduced it onto the market – whereas classifying it as an autonomous subject entails attributing liability to the robot itself, since its autonomous subjectivity and decision-making capacity would imply the assumption of corresponding responsibilities[20].

In the latter case, however, the problem of the robot's actual solvency and the injured party's real and effective possibility of obtaining compensation remains unresolved.

It is therefore necessary to reconsider how robots are conceptualised, the type of liability to be attributed to them for the damages they cause, and the economic coverage of compensation.

The solution of attributing electronic personality to robots, thus considering them as a tertium genus distinct from natural and legal persons, may be viable in light of their growing decisional and behavioural

[20] Al riguardo, U. Ruffolo, *Le responsabilità da intelligenza artificiale e le questioni di libertà d'espressione e copyright,* cit., p. 490; V. Sanasi D'Arpe, *cit.*, p. 15.

autonomy, which endows them with an individuality approaching that of human beings[21].

This appeared to be the direction taken by the European Parliament, which, in its Resolution of 16 February 2017, proposed the introduction of mandatory insurance coverage for damages caused by robots as a means of ensuring compensation for harm suffered by third parties and directly attributable to robots rather than to the human actors who produced or marketed them (point 56). Similarly, in its Resolution of 12 February 2019, the Parliament expressly called on the Commission to promote a regulatory environment conducive to the development of Artificial Intelligence and to regularly review existing legislation to ensure its suitability for AI while respecting the EU's fundamental values (point 114).

For all the reasons outlined above, there is no doubt that a new and more specific regime of civil liability for robots and Artificial Intelligence systems is required, capable of better capturing and regulating the phenomenon in all its specificities. Nevertheless, pending the hoped-for legislative intervention, the application of producer liability rules remains preferable, with a view to enhancing victim protection and achieving a more equitable distribution of the costs, risks, and benefits associated with the use of robots and Artificial Intelligence systems technoregulation).

2. Prospects for Reform in Balancing Technological Development and the Protection of Fundamental Rights

The European legislator has followed the development of Artificial Intelligence systems and the effects that such development has produced on the market and therefore intervened, on 23 October 2024, with the adoption of Directive (EU) 2024/2853 on liability for defective products (which repeals Directive 85/374/EEC, with effect from 9 December 2026, and will apply to products placed on the market or put into service after that date).

The new regime on liability for damage caused by defective products overcomes interpretative doubts as to the definition of a defective product[22]

[21] See G. Sartor, *Gli agenti software: nuovi soggetti del ciberdiritto?,* in «Contr. impr.», II (2002), p. 465; Id., *Gli agenti software e la disciplina giuridica degli strumenti cognitivi,* in Il diritto dell'informazione e dell'informatica, III (2003), pp. 55 ff.; C. Boscarato, *Who is responsible for a robot's actions? An initial examination of Italian law within a European perspective,* in van den Berg, Klaming (eds.), Technologies on the Stand: Legal and Ethical Questions in Neurosciences and Robotics, Nijmegen, 2011, pp. 93 ff.; D.C. Vladeck, *Machines without principals: liability rules and artificial intelligence,* in «Washington Law Review», Vol. 89, CXVII (2014).

[22] ECJ, 10.06.2021, C-65/20.

and the inclusion of intelligent products within the scope of that notion, which will therefore also take into account the product's ability to learn or acquire new functionalities even after it has been placed on the market[23]. It thus seeks to ensure that, in the event of damages caused by a defective product, the victim may rely on the same level of protection regardless of whether the damages was caused by a digital or a material product.

Strict liability is regarded as the only appropriate solution to address the problem of the fair allocation of risk arising from technological products[24], including software, even where integrated into other movable goods or incorporated into immovable property[25].

Technological development offers products that may be tangible or intangible, and software – such as operating systems, firmware, and Artificial Intelligence applications – is increasingly widespread on the market. Software may be placed on the market as a standalone product or integrated into other products as a component. Software must be considered a product, unlike information, that is the content of digital files, such as multimedia files or software source code[26].

The directive's regime also applies to digital services, thereby overcoming the traditional distinction between product and service, since the digital service is typically integrated into products or interconnected with them and therefore contributes to product safety in the same way as physical or digital components. Consequently, digital services are considered components of the product into which they are integrated or with which they are interconnected where they are under the control of the product manufacturer[27]

Accordingly, the directive includes within the definition of "product" digital manufacturing files, raw materials, and software[28].

The technological character of a product therefore no longer negatively affects the applicability, in the case at issue, of the more victim-friendly regime of the producer's strict liability.

The victim-favourable regime is further strengthened by the presumption of defectiveness where the alleged injuring party fails to disclose relevant evidence, where the product does not comply with mandatory safety requirements laid down by EU or national law, or where the damages was caused by an evident malfunction of the product during reasonably forese-

[23] Art. 7, co. 1, of directive.
[24] Recital n. 2 of directive.
[25] Recital n. 6 of directive.
[26] Recital n. 13 of directive.
[27] Recital n. 17 of directive.
[28] Art. 4 of directive.

eable use or under ordinary circumstances; and by the presumption of causation between the product's defectiveness and the damage where it has been proven that the product is defective and that the nature of the damage is consistent with the alleged defect[29]. Moreover, the Court may presume the defectiveness of the product, the causal link between the defect and the damage, or both, where the claimant encounters excessive difficulties, due to technical or scientific complexity, in proving defectiveness or causation, or where the claimant demonstrates the likelihood that the product is defective or that such a causal link exists[30].

By introducing a presumption of causation, the legislator intended to incorporate into the legal system an institution that will need to be defined in terms of its content and practical configuration. It nevertheless remains the case that the use of presumptions facilitates the injured party's burden of proof and significantly contributes to mitigating the risk associated with the product's technological nature.

In any event, account should be taken of the issues posed by the so-called Internet of Things (IoT) and by development risk, both of which are closely linked to the phenomenon of the intelligent product.

Indeed, the latter may have been manufactured and placed on the market without any original defect, yet may develop a defect as a result of its ability to learn from its surrounding environment and through interaction with other intelligent products.

The harmful event may therefore arise from the relationship between the product at issue and other products attributable to different manufacturers, making it more difficult to identify the defect that caused the damage and, consequently, to attribute liability to one manufacturer rather than another.

With regard to development risk, the fact that an intelligent product may acquire a defect as a result of interaction with other intelligent products attributable to different manufacturers, or due to its learning capability, requires a reassessment of the traditional notion of product defect so as to include supervening, non-original defects. Nevertheless, Article 11 of the directive excludes the liability of the economic operator where it is proven that it is likely that the defect that caused the damage arose after the product was placed on the market, essentially assuming that once the product is placed on the market the manufacturer loses control over it. In reality, the manufacturer's control over the product is a

[29] Art. 10, co. 2 e 3, of directive.

[30] Art. 10, co. 4, of directive.

decisive criterion for determining whether a product is defective, and digital technologies enable manufacturers to exercise control over products even after they have been placed on the market. Consequently, the manufacturer remains liable for defects arising after that moment and caused by software or related services under its control, for example as a result of updates, improvements, or machine-learning algorithms[31]. Furthermore, Recital 32 of the directive opens the way to an assessment of product safety that also takes into account the moment when the product leaves the manufacturer's control.

The purpose of the new legislation is therefore to maximise protection for those harmed by Artificial Intelligence tools, even though it primarily concerns high-risk Artificial Intelligence systems and pursues only minimal harmonisation.

Finally, on 13 June 2024, the European Parliament adopted the Artificial Intelligence Act ("AI Act")[32], a measure aimed at comprehensively regulating Artificial Intelligence systems in light of, and with regard to, the rights and values of the European Union. By virtue of that regulation, a uniform legal framework has been established for the development, placing on the market, putting into service, and use of Artificial Intelligence systems and applications, including and especially forms of generative Artificial Intelligence, which represent the new frontier of technology.

The aim of the AI Act is to improve the functioning of the internal market and promote the deployment of human-centric and trustworthy Artificial Intelligence, while ensuring a high level of protection of health, safety, and fundamental rights enshrined in the Charter of Fundamental Rights of the European Union, including democracy, the rule of law, and environmental protection, against the harmful effects of Artificial Intelligence systems in the European Union, while supporting innovation[33].

The regulation lays down harmonised rules for the placing on the market, putting into service, and use of Artificial Intelligence systems in the Union; prohibits certain Artificial Intelligence practices; sets specific requirements for high-risk Artificial Intelligence systems and obligations for operators of such systems; establishes harmonised transparency rules for certain Artificial Intelligence systems; lays down harmonised rules for the placing on the market

[31] Recital n. 50 of directive.

[32] Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024, laying down harmonised rules on Artificial Intelligence and amending Regulations (EC) No. 300/2008, (EU) No. 167/2013, (EU) No. 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act).

[33] Recital n. 1 of Regulation (EU) 2024/1689. Cassano, Tripodi (eds.), Il Regolamento europeo sull'Intelligenza Artificiale. Commento al Reg. UE N. 1689/2024, Santarcangelo di Romagna, 2024.

of general-purpose Artificial Intelligence models; and provides rules on market monitoring, market surveillance, governance, and enforcement, as well as measures to support innovation[34].

The Artificial Intelligence Act is based on a risk-based approach and therefore regulates the phenomenon by distinguishing different levels of risk, depending on the impact and incidence that technology may have on the individual's existence and fundamental rights.

Accordingly, the risk-based approach is the framework used by the EU legislator to regulate the use of Artificial Intelligence and to differentiate the related obligations depending on the risk that Artificial Intelligence machines and applications may pose to the individual's fundamental rights.

Manufacturers, providers, importers and distributors and, more generally, users of Artificial Intelligence systems are required to take account of the relevance of such risks, assuming the corresponding responsibility that may arise where an intelligent product causes damages to third parties.

Protection is ensured through Article 99 of the Regulation, which refers to Member States the task of establishing rules on sanctions and other enforcement measures intended to be effective, proportionate, and dissuasive. This does not, however, preclude the attribution of liability according to the criteria and principles outlined above, especially in the light of pending legislative proposals.

Guido Alpa therefore concludes by proposing the establishment of an articulated system which, alongside civil liability rules, also includes private insurance and a public guarantee fund, so that risk is initially transferred to insurance and, where the victim is not fully compensated, to the fund[35].

Interpretative uncertainties remain numerous, and it is necessary to continue researching models of civil liability capable of satisfying the primary need to protect the fundamental rights of the human person while, at the same time, taking account of their economic impact.

The law must regulate the phenomenon of new technologies without, however, limiting – or worse, obstructing – their development.

[34] Art. 1 of Regulation (EU) 2024/1689.

[35] G. Alpa, cit., p.139; V. Sanasi D'Arpe, cit., pp. 44 ff.

BIBLIOGRAPHY

Alesso D., *Robot e A.I.: questioni di soggettività giuridica*, in cyberlaws.it, (2018).

Alpa G., *L'intelligenza artificiale. Il contesto giuridico*, Mucchi Editore, Modena, 2021.

Amato C., *Responsabilità civile da "intelligenza" artificiale: prospettive di disciplina europea*, in «Rivista trimestrale di diritto e procedura civile», (2025).

Bayern S., *The Implications of Modern Business-Entity Law for the Regulation of Autonomous Systems*, in «Stanford Technology Law Review», Vol. 19, XCIII (2015), pp. 93 ff.

Bertolini A., *Robots as Products: The Case for a Realistic Analysis of Robotic Applications and Liability Rules*, in «Law, Innovation and Technology», Vol. 5, II (2013), pp. 214 ff.

C. Boscarato, *Who is responsible for a robot's actions? An initial examination of Italian law within a European perspective*, in van den Berg, Klaming (eds.), Technologies on the Stand: Legal and Ethical Questions in Neurosciences and Robotics, Nijmegen, 2011, pp. 93 ff.

Gaeta M.C., *Automazione e responsabilità civile automobilistica*, in «Responsabilità civile e previdenza», V (2016), pp. 1718 ff.

Gentili A., *Regole per l'intelligenza artificiale*, in «Contratto e impresa», Vol. 40, IV (2024), pp. 1043 ff.

Geistfeld M., *A Roadmap for Autonomous Vehicles: State Tort Liability, Automobile Insurance, and Federal Safety Regulation*, in «California Law Review», Vol. 105, VI (2017), pp. 1611 ff.

Holder C., Khurana V., Harrison F., Jacobs L., *Robotics and law: key legal and regulatory implications of the robotics age*, in «Computer Law & Security Review», Vol. 32, III (2016), pp. 383 ff.

Iurilli C., *Intelligenza artificiale, tra essere e res. Alla ricerca della responsabilità extracontrattuale o di una efficace garanzia patrimoniale*, in «Judicium», (2025).

Karnow C.E.A., *The application of traditional tort theory to embodied machine intelligence*, in R. Calo, A.M. Froomkin, I. Kerr (eds.), *Robot Law*, Edward Elgar Publishing, Cheltenham, 2016.

Leenes R.E., Palmerini E., Koops B., Bertolini A., Salvini P., Lucivero F., *Regulatory challenges of robotics: some guidelines for addressing legal and ethical issues*, in «Law, Innovation and Technology», IX (2017), pp. 1 ff.

Matthias A., *The responsibility gap: Ascribing responsibility for the actions of learning automata*, in «Ethics and Information Technology», VI (2004), pp. 175 ss.

Novelli C., Casolari F., Hacker P., Spedicato G., Floridi L., *Generative AI in EU Law: Liability, Privacy, Intellectual Property, and Cybersecurity*, in «Computer Law & Security», (2024).

Palmerini E., *Il problema dei (nuovi) danni negli illeciti algoritmici*, in «Responsabilità civile e previdenza», (2025).

Pardolesi R., Davola A., *In viaggio col robot: verso nuovi orizzonti della r.c. auto (driverless)?,* in «Danno e resp.», (2017), pp. 616 ff.

Procida Mirabelli di Lauro A., *L'ingiustizia del danno nel sistema della responsabilità civile*, in «Danno e responsabilità», IV (2025), pp. 409 ff.

Proietti G., *Definire l'indefinibile? I sistemi di intelligenza artificiale alla ricerca di un inquadramento sistematico*, in «Contratto e impresa», XII (2024), pp. 882 ff.

Rabin R.L., Abraham K.S., *Automated Vehicles and Manufacturer Responsibility for Accidents: A New Legal Regime for a New Era*, in «Virginia Law Review», Vol. 105, CXXVII (2019), pp. 127 ff.

Ruffolo U., *Le responsabilità da intelligenza artificiale e le questioni di libertà d'espressione e copyright*, in «Contratto e impresa», II (2025), pp. 489 ff.

Ruffolo U., *Per i fondamenti di un diritto della robotica self-learning; dalla machinery produttiva all'auto driverless: verso una "responsabilità da algoritmo"?,* in U. Ruffolo (ed.), *Intelligenza artificiale e responsabilità*, Milano, Giuffrè, 2018, pp. 1 ff.

Ruffolo U., *Self-driving car, auto driverless e responsabilità*, in U. Ruffolo (a cura di), *Intelligenza artificiale e responsabilità*, Milano, 2018.

Rumi T., *La responsabilità del produttore nell'era dell'A.I.*, in «Danno e responsabilità», (2025), pp. 436 ff.

Sanasi d'Arpe V., *Intelligenza artificiale: le questioni della soggettività giuridica, della responsabilità e dell'imputazione delle condotte*, in «Iurisdictio. Storia e prospettive della giustizia», V (2024), p. 6.

Santosuosso A., Boscarato C., Caroleo F., *Robot e diritto: una prima ricognizione*, in «Nuova giurisprudenza civile commentata», VIII-VIII (2012), pp. 1 ff.

Sartor G., *Gli agenti software: nuovi soggetti del ciberdiritto?*, in «Contratto e impresa», II (2002), pp. 465 ff.

Sartor G., *Gli agenti software e la disciplina giuridica degli strumenti cognitivi*, in «Il diritto dell'informazione e dell'informatica», (2003), pp. 55 ff.

Scognamiglio C., *Responsabilità civile ed intelligenza artificiale: quali soluzioni per quali problemi?*, in «Responsabilità civile e previdenza», IV (2023), pp. 1073 ff.

Smith B.W., *Automated Driving and Product Liability*, in «Michigan State Law Review», I (2017).

Sunghyo K., *Crashed Software: Assessing Product Liability for Software Defects in Automated Vehicles*, in «Duke Law & Technology Review», XVI (2018), pp. 300 ff.

Teubner G., *Rights of non-humans? Electronic agents and animals as new actors in politics and law*, in «Journal of Law and Society», Vol. 33, (2006).

Vladeck D.C., *Machines without principals: liability rules and artificial intelligence*, in «Washington Law Review», Vol. 89, CXVII (2014).

Woodward J.E., *Oops, My GPS Made Me Do It! GPS Manufacturer Liability under a Strict Product Liability Paradigm*, in «University of Dayton Law Review», Vol 34, III (2009), 429 ff.

FERNANDA FAINI

Artificial Intelligence and Cybersecurity: European Law and Regulatory Tools

Abstract

The paper aims to reflect on the relationship between law and technology, examining the face of digital society, in order to analyse the current configuration of European legal regulation in this area, the peculiarities that characterise it and the directions that emerge for the future. The contribution examines the philosophical, ethical and legal model of governance of technology and, in particular, Artificial Intelligence, as it emerges in the light of the European legislation that has been taking shape in recent years, aimed at regulating the digital dimension in its various aspects (*Artificial Intelligence Act, Data Governance Act, Data Act, Digital Services Act, Digital Markets Act*, etc.), also paying attention to the changing geometries of power. European regulation denotes an evolution of law both in terms of the form of legal regulation and in terms of instruments provided by the legislature, which innovate traditional paradigms.

Il contributo si propone di riflettere sul rapporto tra diritto e tecnologia, esaminando il volto della società digitale, al fine di analizzare l'attuale configurazione della regolazione giuridica europea in tale ambito, le peculiarità che la contraddistinguono e le direttrici evolutive che si delineano per il futuro. Lo studio esamina il modello filosofico, etico e giuridico di governance della tecnologia e, in particolare, dell'Intelligenza Artificiale, così come emerge alla

* Associate Professor of Legal Informatics and Digital Law (Philosophy of Law) at the Department of Law, Pegaso University; Deputy Director of the Interuniversity Research Centre for Law and Ethics of Innovation, Technology and Artificial Intelligence (LEITAI), a joint initiative of the San Raffaele, Pegaso, and Mercatorum Universities.

§ This contribution develops the presentation at the conference 'Cybersecurity e prodotti: sfide e responsabilità nelle società interconnesse', 29 November 2024, Italian Supreme Court.

luce della normativa europea che si è progressivamente delineata negli ultimi anni, volta a disciplinare la dimensione digitale nelle sue molteplici articolazioni (*Artificial Intelligence Act, Data Governance Act, Data Act, Digital Services Act, Digital Markets Act,* etc.), con particolare attenzione anche alle mutevoli geometrie del potere. La regolamentazione europea testimonia un'evoluzione del diritto sia sotto il profilo della forma della regolazione giuridica, sia con riferimento agli strumenti messi a disposizione dal legislatore, che innovano i paradigmi tradizionali.

1. Legal rules in the digital society

Over time, information technologies have managed to overcome their ontological essence as mere tools in the hands of man, becoming a multifaceted means capable of bringing about a profound transformation in human existence itself, accompanied by the emergence of new needs and an impact on human activities and relationships.

In this context, law, the instrument with which humans regulate life, is called upon to regulate information technology and govern the digital dimension in order to protect rights and balance different and sometimes conflicting interests, resolving conflicts. In performing its function, law is exposed to change and is affected by the peculiarities of the object it aims to regulate, namely the digital dimension; as a result, regulation is bound to change as a result of a number of factors.

A first aspect derives from the subject matter of the regulation, from the technology itself, which is accompanied by the presence of an ecosystem of rules that is different from the legal one, namely the rules applied by the computer code (lex informatica or digitalis)[1]. Technologies are dominated by codes, instructions and IT rules, which are capable of influencing human behaviour, since they enable or prevent actions and interactions, enabling or disabling them, defining their methods and constraints, linking effects, determining what information to provide to the user and, consequently, the degree of transparency and comprehensibility for those who use them. The incisive phrase 'Code is Law', used by Lessig, highlights the regulatory aspect inherent in computer code[2].

[1] See L. Lessig, *The Law of the Horse: What Cyberlaw Might Teach*, in Harvard Law Review, Vol. 113, II (1999), pp. 501-546; G. Sartor, *L'informatica giuridica e le tecnologie dell'informazione. Corso di informatica giuridica*, II ed., Giappichelli, Turin, 2010, p. 37 ff.; E. Maestri, *Lex informatica e diritto. Pratiche sociali, sovranità e fonti nel cyberspazio,* in «Ars Interpretandi», I (2017), pp. 15-28.

[2] L. Lessig, *Code and Other Law of Cyberspace,* Basic Books, New York, 1999.

Therefore, in determining what is technologically possible, lex informatica has the capacity to influence all other forms of regulation, including legal regulation, and humans, through the law, must be able to govern it. To do so, it is necessary to achieve a difficult balance based on flexibility and adaptability, capable of ensuring the effectiveness of legal rules without limiting the potential of technological evolution, but at the same time firmly grounded in predictability and legal certainty, without allowing technology to dominate regulation[3].

This raises ontological and legal issues closely related to the technical characteristics that dominate the digital scene and, in particular, one of its protagonists, Artificial Intelligence, such as the mechanism of inferences and correlations on which algorithms are based, the opacity and related difficulty of understanding by humans (black box) in machine learning and deep learning, the related difficulty of motivating the results obtained by the machine[4], as well as potential errors and biases. Algorithms can lead to technical problems, but also to related ethical, social and legal issues due to the presence of errors and poor-quality data, biases and prejudices, and potential human manipulation. As a result, they can be a harbinger of discrimination, inequality and disparity[5].

Technology is only theoretically 'neutral', since in its instrumental value it acquires the meaning that humans give it when they use it, despite the fact that the law pursues its neutrality in the relevant acts[6]; technologies are 'double-edged swords' according to Bauman's effective configuration[7]. Consequently, the asymmetries and disparities typical of society can emerge in the digital context: the digital dimension can resemble the analogue one in its dark side.

This first aspect characterising the relationship between law and technology is linked to a second aspect: in the digital dimension, the object of regulation changes, consisting of intangible assets, data, which are different from the res corporales to which the law is traditionally accustomed. Objects and subjects become digital; assets are transformed into services; the entrenched paradigm of

[3] V. Frosini, *Il diritto nella società tecnologica,* Giuffrè, Milan, 1981.

[4] *Infra*, § 3.3.

[5] G. Corasiniti, *Tecnologie intelligenti. Rischi e regole,* Mondadori, Milan, 2023, p. 63 ff.

[6] According to the principle of technological neutrality, legal effects must be 'attainable by any technical means, provided that the requirements' laid down are met (recital 27, EU Regulation No 910/2014). The principle is linked to the need to protect competition, respect autonomy, ensure effectiveness and allow adaptation to technological developments; G. Finocchiaro, *Diritto di Internet,* III ed., Zanichelli, Bologna, 2020, p. 88 ff.

[7] Z. Bauman, D. Lyon, *The Sixth Power: Surveillance in the Age of Liquid Modernity,* Italian translation, GLF Editori Laterza, Rome-Bari, 2015, p. 86 ff.

ownership is undermined by the paradigm of access to data, services, and one's digital existence[8]. As a result, data and technologies capable of extracting value from it, such as Artificial Intelligence, whose fundamental elements lie precisely in data (the 'soul') and algorithms (the 'engine'), are taking on a dominant role[9]. In the digital society, well-being and human development depend significantly on data-based technologies; today's society, as a *data society*, is intimately pervaded by data, ultimately shaping humans themselves as *data subjects*[10].

Linked to the centrality of data and the network is another particularly significant aspect that characterises contemporary society dominated by emerging technologies; the development of the Internet and Web 2.0 outlines a model of equal, horizontal and collaborative production based on dissemination, sharing and reuse, which sits alongside established vertical models based on the dominant paradigm and the control of a central authority[11]. The paradigms of ownership and control, characteristic of traditional law, are giving way to a leaderless, transnational and more egalitarian approach, which also has an impact on the reference space and on those who actually produce rules.

The dimension of space is also affected by the digital revolution and, as a result, influences law; the field of action is changing as territorial boundaries are overcome. Different spatial coordinates are emerging, due to the global dimension of reference, which has no territorial or 'sovereign' boundaries, determining the relevance of supranational law in this area.

The digital space/non-space is leading to a redefinition of the public and private spheres, and a new power structure is emerging, in which states are weakened and their power over the digital world is eroded due to the global scope of action and the presence of technological giants, 'controllers of the toll for access to digital life', favoured by the supranational dimension itself, by their ability to use technology, by freedom of trade, by the wide net or by the absence of provisions appropriate to the changed context. Tech giants have been able to dominate the global dimension thanks to the tools of private contractual autonomy through the concrete regulation of access to digital services, thus affecting rights[12].

[8] See J. Rifkin, *L'era dell'accesso. La rivoluzione della new economy,* Italian translation, Mondadori, Milan, 2000; L. Floridi, *La rivoluzione dell'informazione,* Codice edizioni, Turin, 2012, p. 15; S. Rodotà, Il diritto di avere diritti, Laterza, Roma-Bari, 2012, p. 105 ff.

[9] The 'soul' of Artificial Intelligence consists of enormous volumes of data, known as big data, which comes from various sources and is analysed using algorithms, *data mining, big data analytics, machine learning, deep learning and* other techniques.

[10] On this subject, see F. Faini, *Data society. Governo dei dati e tutela dei diritti nell'era digitale,* Giuffrè Francis Lefebvre, Milan, 2019.

[11] See G. Sartor, *Internet e il diritto,* in C. Di Cocco, G. Sartor (eds.), *Temi di diritto dell'informatica*, 4th ed., Giappichelli, Turin, 2020, p. 1 ff.

[12] See C. Accoto, *Il mondo dato. Cinque brevi riflessioni di filosofia digitale,* Egea, Milan, 2017, p. 108.

As a result, the entities that actually produce rules are changing in a context of disintermediation, in which new asymmetries and power imbalances are emerging: Big Tech companies, such as Amazon, Apple, Facebook, Google and Microsoft, are creating an algorithmic order in which they act as legislative (dictating the rules), executive (enforcing them) and judicial (judging violations with sanctions such as account suspension or deactivation) powers. These rules are widely accepted, albeit more or less consciously, and are capable of influencing individuals' lives as much as the binding rules produced by states. The new 'territories' constituted by digital platforms are perceived by the community as public legal spaces, although they lack democratic legitimacy and are guided not by the public interest but by economic profit[13].

In this context, where the law is affected by the peculiarities of the subject matter it regulates and must contend with changing power dynamics, a comprehensive legal framework has taken shape at European level in recent years to regulate the digital dimension of human life. Initially consisting mainly of soft law instruments, it has more recently been supplemented by a set of regulations such as the Cybersecurity Act (EU Regulation 2019/88) Data Governance Act (EU Regulation 2022/868), the Digital Markets Act (EU Regulation 2022/1925), the Digital Services Act (EU Regulation 2022/2065), the *Data Act* (EU Regulation 2023/2854) and the *Artificial Intelligence Act* (EU Regulation 2024/1689; hereinafter also referred to as the AI Act).

This contribution aims to analyse the current configuration of European legal regulation in this area, its distinctive features and the directions emerging for the future. Specifically, the contribution aims to examine the philosophical, ethical and legal model of technology governance and, in particular, Artificial Intelligence and related aspects of Cybersecurity, which emerges in light of the aforementioned European regulation of recent years, aimed at regulating the digital dimension in its various aspects, while also paying attention to the changing power geometries.

The European regulation in fact, denotes an evolution of law both in terms of form and the 'mode of expression' of law[14], and in terms of the substance of the solutions and instruments provided by the legislator,

[13] With regard to the classification of digital platforms as public legal spaces, please refer to the extensive literature on the subject; see, inter alia, G. De Gregorio, *Digital Constitutionalism in Europe. Reframing Rights and Powers in the Algorithmic Society,* Cambridge University Press, Cambridge, 2022; H. Micklitz, O. Pollicino, A. Reichman, A. Simoncini, G. Sartor, G. De Gregorio (eds.), *Constitutional Challenges in the Algorithmic Society*, Cambridge University Press, Cambridge, 2021.

[14] *See below*, § 2.

which innovate traditional paradigms[15], partly as a result of the fact that legal regulation in the digital dimension has often proved imperfect in practice, with the protection of human rights also often proving imperfect.

2. The new face of law: changes in the form of regulation

The law must succeed in the complex challenge of being sustainable in the digital dimension and, to this end, it is called upon to achieve a wise balance between different rights and interests. In light of the characteristics of the digital reality, this is only possible with effective governance and a transnational vision of the problems, accompanied by the related need to ensure consistency in the responses offered by different legal systems, so as not to generate discrepancies between the rule and what it is intended to regulate.

From the first perspective, the need to build effective governance of the digital world has been apparent since the 1990s with the establishment of independent bodies responsible for protecting specific rights and interests, such as, at national level, the Authority for Information Technology in Public Administration (AIPA), currently the Agency for Digital Italy (AgID), the Data Protection Authority, the Communications Regulatory Authority (Agcom) and the Competition and Market Authority (AGCM). In the concrete and uniform application of legal provisions and in balancing different rights and interests, the role of the authorities is also fundamental in order to respond to the aforementioned demands for flexibility and adaptability, which are necessary for legal rules to be effective in the digital dimension and capable of addressing technological developments through the interpretation of the relevant principles and standards.

The need to ensure effective governance is also reflected in the new European regulations, which provide for the establishment of committees and expert groups at supranational level, dedicated to specific areas of action or technologies (Artificial Intelligence, data, digital services, digital markets) and called upon to assist and advise the European Commission in the uniform application of regulations, such as *the European Artificial Intelligence Board*[16]*, the European Data Innovation Board*[17]*,*

[15] *See below*, § 3.

[16] Art. 65 et seq., Title VI, EU Regulation 2024/1689 *AI Act,* which, in addition to *the Board, provides for the European AI Office, the Advisory Forum* and a Scientific Committee composed of independent experts in *the governance* of Artificial Intelligence.

[17] Art. 29 et seq., EU Reg. 2022/868 *Data Governance Act.*

the European Board for Digital Services[18]*, the high-level group for the Digital Markets Act*[19]*.*

This need for governance highlights the need to reclaim a public legal space to protect individual rights, given the imbalance in the digital sphere caused by the power wielded by private entities such as Big Tech companies. However, the risk lies in tipping the balance towards one or the other right to be protected or towards the technology for which the specific authority is established and consequently acts, losing the holistic approach needed to address the complexity of the digital reality, characterised by the presence of different rights and interests and the need to find a complex balance and a difficult equilibrium between them. The provision for supranational bodies to contribute to the effective and uniform application of the rules is a commendable choice in the effort to identify independent entities dedicated to governing the digital dimension, it also necessitates synergistic cooperation between the various bodies, going beyond the sectoral scope of action that would otherwise risk losing the holistic approach required in light of the obvious intersections between the various regulations. For example, when it comes to Artificial Intelligence, *data governance* will also come into play, even though it is subject to separate regulations and different bodies, which will therefore necessarily have to cooperate with each other.

The need for a holistic approach to addressing the digital dimension also underpins the second aspect (the transnational logic), since the level at which we need to look is necessarily supranational: the global reality of the network implies a change in the geographical boundaries of regulation and makes it necessary to arrive at shared solutions on the protection of rights. The supranational dimension is necessary to ensure the effectiveness of the rules, avoiding the otherwise inevitable tension between the global dimension of the issues and the territorial dimension of the provisions to be applied. In this respect, the supranational impulse of the European legislator has been evident since the 1990s, pushing in the new millennium, in parallel with the greater pervasiveness of technology in human life, for the systematic and organic regulation of entire areas of law.

In addition to the extension of the territorial scope, the European reference regulation also highlights the breadth of the topics covered, which is combined with the heterogeneity of the acts containing the rules: not only *hard law* in the strict sense, but also soft law, policy documents,

[18] Art. 61 et seq., EU Reg. 2022/2065 *Digital Services Act.*

[19] Art. 40, EU Reg. 2022/1925 *Digital Markets Act.*

guidelines and strategies of institutional bodies, independent authorities, organisations, committees, and even task forces and expert groups; the complexity of the phenomena is addressed with heterogeneous instruments produced not only by the competent public institutions, but also by private entities[20].

In order to regulate technology effectively, the law must take into account and integrate different ecosystems of rules, such as IT rules, which dictate 'law' in the digital dimension, and those established in the sphere of private autonomy by tech giants, which regulate the 'territories' constituted by the extensive digital platforms over which they exercise their dominance. Consequently, regulation requires a multilevel genesis, a multistakeholder approach and the integration of heterogeneous sources in order to make use of the various stakeholders and the resulting shared responsibility of the different producers of rules, whether legal or not, in our contemporary world: public powers, but also private powers. Therefore, alongside the rules laid down by public authorities, there are forms of co-regulation and self-regulation, where the rules are set by the recipients themselves, as in the case of Big Tech companies and a sort of contemporary lex mercatoria; states are called upon to engage in new forms of cooperation and collaboration.

The heterogeneity of the types of acts derives precisely from the characteristics of the object of regulation: technology is characterised by technical specificity, scientific progress and constant evolution, which require qualified expertise and entai the need for precise acts, technical rules and flexible instruments to deal with it, while the norm in the strict ontological sense is characterised by being general and abstract, the result of a long process stemming from the democratic process and intended to last over time. Consequently, a complex set of different rules is needed in this area, capable of establishing a comprehensive and sustainable (and therefore effective) regulatory framework and, to this end, aimed at ensuring flexibility and adaptability to technological developments and changes, but also capable of ensuring predictability and certainty, which are fundamental principles of law[21] .

The European regulatory framework reflects this ongoing tension between flexibility and predictability, with the aim of ensuring effectiveness but also certainty, protecting individuals and striking a virtuous balance between different interests.

[20] A. Simoncini, *Verso la regolamentazione della Intelligenza Artificiale. Dimensioni e governo,* in *BioLaw Journal,* 2, 2021, p. 412.

[21] See C. Casonato, B. Marchetti, *Prime osservazioni sulla proposta di regolamento dell'Unione Europea in materia di intelligenza artificiale*, in «BioLaw Journal», III (2021), pp. 417-418.

The need to ensure flexibility and adaptability in the face of relentless technological development and the resulting need for updating calls for flexibility mechanisms, such as the general obligation in the *Artificial Intelligence Act* to review the regulation and the procedure for reviewing and amending the annexes relating to the identification of high-risk systems, even outside the ordinary legislative procedure that would otherwise be necessary. in the same vein, the provisions relating to the *regulatory sandbox* mechanism, which are useful for adapting provisions to a particularly complex reality[22] , and the use of codes of good practice and codes of conduct, which contribute to the correct application of the rules[23] , can be interpreted.

The need to ensure predictability and legal certainty, together with the need to guarantee uniformity of solutions, have led to a change in the European Union's approach, with a tendency in recent years to 'tighten' the force of the legislative acts adopted (from *soft law* to *hard law*) in order to ensure their effectiveness. This is the case with the European regulation, which, with its greater influence on Member States in recent years, is replacing the directive. The European Union is creating a legal framework consisting of a set of regulations aimed at regulating the digital dimension (*AI Act, Cybersecurity Act, Data Governance Act, Data Act, Digital Services Act, Digital Markets Act*)[24]. In this regard, it should be noted that the heterogeneity of acts is not being abandoned in order to ensure the effectiveness of the rules laid down: these provisions also refer to *soft law* acts necessary for the application of the rules, such as the aforementioned codes of conduct. Furthermore, although the formal instrument chosen to regulate the digital dimension is the regulation, in addition to the flexibility mechanisms examined above, the various acts leave ample room for application by Member States, which balance the greater force of the act with the need to ensure adaptability and effectiveness, even if this may also cause uncertainty and potential discrepancies in implementation[25] .

[22] The activation of *regulatory sandboxes* allows Member States, for a limited period of time and under the supervision of the competent national authorities, to develop and test innovative Artificial Intelligence systems with a view to their subsequent placing on the market; Art. 57 et seq., Title V, EU Reg. 2024/1689 *AI Act.*

[23] Art. 56 et seq. and Art. 95 et seq., EU Reg. 2024/1689 *AI Act.*

[24] This trend towards abandoning directives in favour of regulations had already been underway for several years, as seen in European Regulations 2016/679 on personal data and 2018/1807 on non-personal data.

[25] In the *AI Act*, for example, the identification of systems posing an unacceptable risk is characterised by the presence of vague and interpretable concepts, which implies flexibility of application and room for manoeuvre in favour of States, but also possible uncertainty in application and related potential discrepancies in implementation.

3. The innovative content of the law: paradigms, instruments and solutions

Alongside these developments in the form of law, the European regulatory framework of recent years contains instruments and solutions that point to future directions and constitute new paradigms of substantive regulation in this area, allowing us to identify innovative content in the law.

First of all, in the European regulatory framework, in addition to the type of act and the form of regulation, the substantive approach is also changing significantly, since the choice has been made to regulate and influence the power structures by affecting the various actors involved, since the aim is to promote development and competitiveness, but also to protect users more effectively by establishing a framework of rules on transparency and accountability, identifying obligations of diligence and transparency for private companies, as in the case of the *Digital Services Act and the Digital Markets Act.*

The European legislator is aware that, in order to build a solid model of technology governance inspired by the principles of democratic systems, it is necessary to overcome existing problems relating to the relationship between law and technology, which are rooted in a number of aspects, such as the critical issues inherent in the very functioning of technology, data and algorithms; the significant imbalance between the parties involved, in particular the weakness of individuals as users and consumers vis-à-vis tech giants; the opacity of data management processes and algorithms.

The analysis of these critical issues also reveals possible solutions, contained in the European regulatory framework, which are capable of innovating traditional paradigms and minimising risks, rebalancing asymmetries in favour of the community and effectively protecting individuals.

The model of technology governance that emerges from the European regulatory framework leverages a number of new tools, solutions and paradigms capable of profoundly innovating the law, which are analysed together with the problematic aspects that characterise them.

3.1 The evolution of the relationship between law and technology: technological humanism, incorporation of law into technology, and the risk-based approach

A first aspect that emerges from the European legal framework concerns the changing and maturing relationship between law and technology.

The direction taken is based on the construction of a technological

humanism[26] and human-centric governance, capable of placing the individual at the centre, for whose protection an adequate institutional framework and a flexible balance between rights are necessary, in a logic that keeps technology as a tool in the hands of man. This approach is embraced by European regulation and, in particular, by *soft law* and *hard law* acts relating to Artificial Intelligence[27]. A holistic approach is emerging, in which the individual, with his or her identity, acts as a prism whose reflections are constituted by the various rights at stake, which are indivisible, albeit in their conflictual nature.

Human-centric governance, guided by an ethical-legal approach, can be achieved and implemented by exploiting technology itself and the relationship between legal and IT rules; the law can use technology to ensure compliance[28].

In this respect, European regulation embraces a preventive and proactive approach, which takes shape in two synergistic and connected aspects: the incorporation of law into technology and the *risk-based approach*.

In fact, European legislation highlights the preventive incorporation of ethical and legal principles, rules and remedies into technology itself as a paradigm capable of resolving or at least minimising the problems arising from the relationship between law and technology, i.e. legal protection by default and by design, based on accountability, which has been present since European Regulation 2016/679 on the protection of personal data[29]. The law can use technology to ensure compliance, creating a model of 'law in technology' (*Code is Law – Law is Code*)[30]. This is a proactive approach that protects individuals from the design stage, by default and through impact assessment, leveraging the configuration of technological systems and Cybersecurity on the one hand, and the empowerment and awareness of individuals on the other, in

[26] See A. Punzi, *Difettività e giustizia aumentata. L'esperienza giuridica e la sfida dell'umanesimo digitale,* in «Ars Interpretandi», I (2021), p. 113 ff.

[27] Before arriving *at the AI Act*, since 2018 the European Union has dedicated a series of acts to the subject, such as communications, plans, recommendations and guidelines. For an analysis *of the provisions of the AI Act*, see, *inter alia*, E. Maestri (ed.), *Introduzione e note di sintesi al Regolamento (UE) 2024/1689 sull'intelligenza artificiale (testo ufficiale e integrale)*, Jovene, Naples, 2024; G. Pascuzzi, *Il diritto dell'era digitale*, 6th ed., il Mulino, Bologna, 2025, p. 303 ff. On the ethical and legal issues raised by Artificial Intelligence, see, *inter alia*, G. Sartor, *L'intelligenza artificiale e il diritto*, Giappichelli, Turin, 2022; F. Casa, S. Gaetano, G. Pascali (eds.), *Intelligenza artificiale: diritto, etica e democrazia*, il Mulino, Bologna, 2025; G. Finocchiaro, *Diritto dell'intelligenza artificiale*, Zanichelli, Bologna, 2024.

[28] On the relationship between ethics and law in the field of Artificial Intelligence, see, inter alia, F.H. Llano-Alonso, *L'etica dell'intelligenza artificiale nel quadro giuridico dell'Unione europea,* in «Ragion Pratica», II (2021), p. 327 ff.; M. Catanzariti, *Etica "artificiale": un nuovo modello regolatorio?,* in «Ars Interpretandi», I (2021), p. 165 ff.

[29] These are the principles of *data protection by design and by default,* accompanied by *data protection impact assessment* (Articles 25 and 35, EU Regulation 2016/679).

[30] In this regard, please refer to F. Faini, *Il diritto nella tecnica: tecnologie emergenti e nuove forme di regolazione,* in *federalismi.it*, XVI (2020), pp. 79-117.

order to confine repression mainly to subsequent sanctions, without unduly restricting the free movement of data and economic development.

This perspective is further developed in the *Artificial Intelligence Act,* which is based on a proactive, risk-based approach, supported by a related system of sanctions, and on its preventive categorisation, distinguishing four categories (unacceptable, high, low and minimal risk), to which differentiated regulation is linked.

The *risk-based approach,* which aims at preventive protection capable of reducing or eliminating the very likelihood of violations occurring, has been embraced by the European Union for its ability to strike a balance between the various interests at stake, since it is aimed at protecting individual rights, but is also geared towards economic growth, since AI producers can consider legal compliance as a production cost that is part of the economic cost of their business rather than facing the risk of having to answer for possible violations[31]. The liability system distinguishes between categories of operators of high-risk and non-high-risk AI systems, justifying the operator's liability by the fact that they are controlling a risk associated with the specific technology. This risk-proportionate approach, which aims to differentiate the regulatory regime in parallel, is aware of the heterogeneity of Artificial Intelligence solutions, in which the autonomy that characterises them, the transparency that can be guaranteed and the human control that can be exercised can vary significantly. The concept of risk also emerges from the *data protection* regulation and the instruments provided for therein.

The proactive approach, which takes shape in the preventive incorporation of law into technology and in the *risk-based approach*, is not without its challenges, which must be addressed in order to ensure the sustainable implementation of the technology governance model. These include, in particular, the ontological 'rigidity' of computer code[32], which conflicts with the flexibility necessary for 'mobile' balancing suitable for the effective protection of rights, as well as linguistic opacity, since the language is computer language and not the natural language of legal norms. It is also complex to identify which ethical and legal principles to incorporate into technologies that are mostly intended for transnational use, given the ethical and legal pluralism that characterises different legal systems. A further significant problem with this approach lies in the fact that compliance with legal principles and the balance between rights

[31] See A. Simoncini, op. cit., p. 413 ff. On this point, reference may also be made to F. Faini, *Intelligenza artificiale e regolazione giuridica: il ruolo del diritto nel rapporto tra uomo e macchina,* in federalismi.it, II (2023), pp. 1-29.

[32] Computer language can lead to the risk of excessive simplification of complex concepts and clauses.

are effectively delegated to those responsible for developing technological solutions and to market operators, with the potential problems that this entails.

3.2 Human in the loop: human supervision and surveillance

The change and evolution of the relationship between law and technology also take shape in a third aspect to be guaranteed in the relationship between humans and machines, namely human supervision and the non-exclusivity of algorithmic decisions made by machines. In this regard, in the regulation dedicated to Artificial Intelligence, the need to keep technology subservient to humans emerges in the attention paid to supervision and human surveillance, aimed at preventing and minimising risks and dangers, in line with the proactive and preventive approach that characterises the technology governance model. In order to build an anthropocentric model, it is necessary to maintain a balance between the human and machine components by ensuring that '*humans in the loop*'. In this regard, it is important to note that technology is no longer just a means of carrying out actions but, increasingly, through *machine* and *deep learning*, it is itself taking significant decisions for humans, where used for such purposes.

To this end*, the Artificial Intelligence Act* provides for concrete measures to ensure substantial human oversight and avoid the risk that human decision-making may be 'pulled' by machine results; the problem, in fact, lies in the influences and 'attractiveness' at a practical level of the solution offered by the machine and in the related difficulty for humans to deviate from their own assessment of what emerges from the technological system, ultimately conforming to the suggested result: this substantially undermines the autonomy of human decision-making[33].

In this regard, however, it is necessary to ask whether and how it is possible to reconcile human supervision and surveillance with *machine* and *deep learning* systems, which are capable of learning autonomously and operate according to decision-making processes that are opaque and unpredictable even for the programmers themselves. It is therefore necessary to verify the actual possibilities for humans to question the results obtained by the machine, check the accuracy of the data and correct the

[33] Art. 14, EU Reg. 2024/1689 AI Act, which provides for measures to enable those responsible for human oversight to understand and interpret the AI system, intervene in the functioning of the system or interrupt it, decide autonomously whether to use it and 'remain aware of the possible tendency to automatically rely on or over-rely on the output produced by a high-risk AI system (automation bias)'. In this regard, see A. Simoncini, *Profili costituzionali della amministrazione algoritmica,* in «Rivista trimestrale di diritto pubblico», IV (2019), p. 1149 ff., according to whom, when a decision-making automatism is introduced into a deliberative process, automatism tends to attract the decision, making it difficult to disregard it and rendering man, who is theoretically its 'master', a 'servant' in practice.

outputs[34]. On closer inspection, explainability and transparency are also necessary conditions for human supervision, since only if humans are able to understand how Artificial Intelligence makes decisions can they exercise effective control and, if necessary, correct its outputs.

3.3 Algorithmic transparency: explainability of Artificial Intelligence and contestability of technology

In synergy with these aspects, in order to overcome the opacity of data management processes and algorithms, it is necessary to ensure enhanced transparency, algorithmic transparency that guarantees knowability, comprehensibility and contestability, making algorithms subject to full knowledge and full scrutiny by humans: this principle translates into ensuring not only information and access to data, but also knowledge of the logic governing algorithms, accompanied by awareness of the consequences and impact on the individual.

With regard to data, algorithms and Artificial Intelligence, therefore, the right to comprehensibility, capable of making the data subject aware, and the right to contestability, enabling the data subject, including through a court, to evaluate and review the decision reached by the technology, must be attributed and recognised. These rights translate more broadly into the right of the individual to maintain autonomy, self-determination and freedom vis-à-vis the machine, which in turn entails the need to understand the mechanisms of operation and a sort of related duty to 'explain' on the part of Artificial Intelligence or, better, on the part of those responsible for it[35]. In line with technological humanism, explainability ensures that Artificial Intelligence remains instrumental to human intelligence and that technology maintains its 'servant' function with respect to humans and their decisions, especially where rights and freedoms are at stake.

At European level, attention to algorithmic transparency has been present since Regulation 2016/679 on data protection: the data controller is required to provide the data subject with the necessary information, including 'the existence of automated decision-making, including profiling [...], and, at least in those cases, meaningful information about the logic involved, as well as the significance and the envisaged consequences of such processing for the data subject'[36]; in the case of automated decision-making, the data subject is granted at least the right to obtain human intervention by the data controller, to express their opinion and to contest the decision, which seems to imply the right to understanding [37].

[34] *See below*, § 3.3.

[35] See M. Palmirani, *Big Data e conoscenza*, in «Rivista di Filosofia del diritto», I (2020), p. 73 ff.

[36] Art. 13, para. 2, letter f), Art. 14, para. 2, letter g) and Art. 15, EU Reg. 2016/679.

[37] Art. 22, para. 3, EU Reg. 2016/679; see U. Pagallo, *Algoritmi e conoscibilità*, in *Rivista di Filosofia del diritto,* 1, 2020, p. 93 ff.

The *Artificial Intelligence Act* pays particular attention to transparency, requiring that clear and adequate information be provided to users in the case of both 'high-risk' and 'low-risk' systems. The European regulation requires that all high-risk systems be designed and developed in such a way as to ensure an appropriate level of transparency (*sufficiently transparent*)[38], thus demonstrating awareness of the need for proportionate measures, without imposing obligations that are unachievable in light of the technological characteristics and therefore destined to be ineffective in practice; the need for sustainability of the law is clearly evident.

Algorithmic transparency must be accompanied by transparency on the part of those who govern algorithms, thus combining the need for machines to be understandable with the need for transparency on the part of the people who manage them. in this regard, the Digital Services Act is relevant, as it imposes transparency and diligence obligations on platforms with regard to the algorithms they use, requiring them to provide explanations to users and assess the systemic risks involved[39].

In this regard, the nature of algorithms raises the question of whether there is always a comprehensible logic, given how they work and the consequent possibility that they may not be intelligible according to logical-rational criteria[40]. Autonomy, in fact, characterises and differentiates Artificial Intelligence from traditional software logic, where the deterministic *'if this, then that'* approach indicates a completely predetermined functioning of the programme, which provides for inputs and outputs: it is precisely in this aspect that the 'intelligent' nature of the machine can be seen, which is capable of learning autonomously through data. In *machine learning*, humans provide the machine with a learning method to apply to the data to which it has access, in order to automatically extract from that data instructions on how to perform the task assigned to it[41]. In *machine learning* and *deep learning* algorithms, Artificial Intelligence itself extrapolates the information necessary to make a decision by analysing large amounts of data and learning from it the numerical parameters needed to make decisions and use them in the subsequent execution phase. the mathematical-numerical representation, i.e. the model generated in the lear-

[38] Art. 13, EU Reg. 2024/1689 *AI Act*.

[39] Articles 4, 34 and 35, EU Regulation 2022/2065.

[40] A. Simoncini, S. Suweis, *Il cambio di paradigma nell'intelligenza artificiale e il suo impatto sul diritto costituzionale,* in «Rivista di Filosofia del diritto», I (2019), p. 97 ff.

[41] In recent years, deep learning techniques have become particularly important. These are based on multi-layered neural networks that reproduce some of the abstract characteristics of the human brain. See F. Lagioia, G. Sartor, *L'intelligenza artificiale per i diritti dei cittadini: il progetto Claudette,* in «Ragion pratica», I (2020), p. 88 ff.

ning or training phase, is not usually directly intelligible to humans (*black box*), an aspect that is particularly relevant from a legal perspective[42].

Therefore, the Artificial Intelligence approach is not based on causal and logical-deductive explanations, typical of human reasoning, but relies on connections and inferences between data and is based on a statistical approach and probability, sometimes leading to difficulties in understanding the reasons (the 'why') behind the answers provided[43]. Therefore, ensuring algorithmic transparency can be particularly complex given the inherent opacity of algorithms, which manifests itself in structural opacity, deriving from their functioning and from the fact that even programmers cannot understand the logical process followed by the machine to arrive at the result starting from the available data, compounded by the linguistic opacity due to computer language rather than natural language[44] .

This aspect can be problematic when Artificial Intelligence is used in legal contexts such as public administration and justice, since both administrative measures (Article 3, Law 241/1990) and judicial measures (Article 111, paragraph 6, Constitution) must necessarily be accompanied by a statement of reasons; the opacity of the reasons prevents any control over the exercise of the function, which could lead to the annulment of decisions[45].

4. Conclusions: philosophical and legal perspectives and scenarios

The prospects outlined above, which emerge from the European framework, make it possible to reduce the problems afflicting the relationship between law and technology, as they address asymmetries and opacity, helping to 'uncover' problems in advance and allowing technological solutions to be designed differently. Furthermore, transparency and accountability allow for *ex post* intervention, where the problem has not emerged *ex ante*, enabling effective protection of the individual.

In light of these considerations, the legal model of technological humanism, which takes shape in European regulation on the subject,

[42] Conditional or deterministic algorithms (if this, then that), on the other hand, are limited to applying predefined computer rules expressed in programming language.

[43] See P. Moro, *Algoritmi e pensiero giuridico. Antinomie e interazioni*, in «Media Law Review», III (2019), p. 19.

[44] On opacity and explainability, see G. Fioriglio, *La Società algoritmica fra opacità e spiegabilità: profili informatico-giuridici,* in «Ars Interpretandi», I (2021), pp. 53 ff.

[45] C. Casonato, *Intelligenza artificiale e diritto costituzionale: prime considerazioni,* in «Comparative and European Public Law», special issue, (2019), p. 119 ff.

makes use of technology, which can implement ethical and legal principles in advance, must be understandable and characterised by substantial human oversight, as well as being based on a *risk-based approach*, governance and accountability on the part of those who govern technology. all these tools are aimed at resolving or at least mitigating the related legal issues.

The European regulatory framework shows an evolution of law, capable of using technology for its own purposes, directing it with an ethical approach towards the values of democratic society: there is a pressing need to recover a public legal space to protect rights and freedoms[46].

In recent years, European regulatory efforts have seen a renewal of law, moving away from a supranational and multistakeholder construct and towards a philosophical-legal approach geared towards the protection of rights. States are called upon to produce rules, recognising the reference provided by a plurality of additional state, supranational or extra-state systems, on the basis of a necessary shared responsibility on the part of the various producers of rules, whether legal or not. In this context, it is appropriate to identify, as provided for in European regulations, supranational committees dedicated to the governance of technology with a certain degree of independence from those powers that may have various interests in steering it in specific directions (public authorities and the market).

After all, the pervasive change brought about by the development of technologies and the central role played by data affects the balance between rights, the balance between interests, the relationship between powers and the upholding of fundamental democratic principles, influencing the ethical and legal guidelines of the future. In this scenario, there is a clear public responsibility to ensure effective protection of individuals, addressing complex balances between rights that intertwine and conflict in the digital dimension, driven by different tensions.

European regulation responds to the need to establish general and abstract principles and rules on the subject and to detail the more technical and flexible aspects in soft law instruments, seeking not to leave rights 'without law', otherwise destined to be protected through interpretation and the supplementary role played by case law and independent authorities, as has often been the case in the past[47].

[46] On Artificial Intelligence and the challenges of constitutionalism, see F. Fabrizzi, L. Durst (eds.), *Controllo e predittività. Le nuove frontiere del costituzionalismo nell'era dell'algoritmo,* Editoriale Scientifica, Naples, 2024.

[47] This is the case of the judgment of the Court of Justice of the European Union of 13 May 2014, case C-131/12, known as *Google Spain*, and, more broadly, of the measures taken by the supervisory authorities, which are crucial in balancing conflicting interests and in the practical application of the rules.

The fundamental philosophical, ethical and legal choices of the European model of technology governance involve the role of the individual, the geometry of power and the relationship between the public and private sectors.

This raises a number of questions. What form of law is necessary for rules to be effective? What model of governance is appropriate to the changed context? What is the *legal horizon of the Internet*[48]?

In this regard, the European regulatory framework and, more generally, the rules governing the relationship between law and technology reveal a tension between certainty/predictability and flexibility/adaptability in order to ensure effectiveness. The challenge for the legal system is to find a middle ground between these demands and a balance between protecting individuals and promoting economic development; there is a need for market regulation in order to prevent abuse by digital companies. It is precisely in order to limit this power, combining the rights of individuals with the interests of the market, that the new European regulations have been approved; after all, it is the task of the law to strike a reasonable balance between different interests.

In this context, there is a need to strengthen the public legal space, which has clearly been lost in the asymmetry between users and platforms. This involves the role of independent authorities, which are ontologically capable of striking that balance between certainty and flexibility and, consequently, creating sustainable law. Consequently, consideration must also focus on the advisability of creating new models of governance and redefining the powers of national authorities and supranational bodies provided for in European regulations, giving rise to joint regulatory enforcement and strong cooperation between authorities[49]. This moves towards a horizontal approach that dismantles vertical regulatory systems, which have always been less suited to the horizontal nature of the network, and avoids the risk of authorities tipping the balance towards the law or technology for which they were established, embracing instead the necessary holistic approach in line with digital complexity and the intertwining of different interests.

Today, it is certainly more necessary than ever to seek what Vittorio Frosini called 'the legal horizon of the Internet', i.e. to identify the '*digital legal system*' and the model of technology governance to which we are and should be directed, which sees a new relationship between law and technology capable of guaranteeing predictability and certainty in human relations, while effectively protecting people from technology with flexible and adaptable to-

[48] See V. Frosini, *L'orizzonte giuridico dell'Internet*, in «Il diritto dell'informazione e dell'informatica», II (2000), p. 271 ff.

[49] *See above*, § 2.

ols: a law that may be imperfect, but is sustainable, capable of strengthening freedoms rather than negatively impacting them thanks to a wise balance between rights and interests, leveraging a holistic approach that places the individual at its centre.

BIBLIOGRAPHY

Accoto C., *Il mondo dato. Cinque brevi riflessioni di filosofia digitale*, Egea, Milan, 2017.

Bauman Z., Lyon D., *Sesto potere. La sorveglianza nella modernità liquida*, GLF Editori Laterza, Rome-Bari, 2015.

Casa F., Gaetano S., Pascali G. (eds.), *Intelligenza artificiale: diritto, etica e democrazia*, il Mulino, Bologna, 2025.

Casonato C., *Intelligenza artificiale e diritto costituzionale: prime considerazioni*, in «Comparative and European Public Law», special issue, (2019), p. 119 ff.

Casonato C., Marchetti B., *Prime osservazioni sulla proposta di regolamento dell'Unione Europea in materia di intelligenza artificiale,* in «BioLaw Journal», III (2021), pp. 1-29.

Catanzariti M., *Etica "artificiale": un nuovo modello regolatorio?*, in «Ars Interpretandi», I (2021), p. 165 ff.

Corasaniti G., *Tecnologie intelligenti. Rischi e regole*, Mondadori, Milan, 2023.

De Gregorio G., *Digital Constitutionalism in Europe. Reframing Rights and Powers in the Algorithmic Society*, Cambridge University Press, Cambridge, 2022.

Fabrizzi F., Durst L. (eds.), *Controllo e predittività. Le nuove frontiere del costituzionalismo nell'era dell'algoritmo*, Editoriale Scientifica, Naples 2024.

Faini F., *Data society. Governo dei dati e tutela dei diritti nell'era digitale*, Giuffrè Francis Lefebvre, Milan, 2019.

Faini F., *Il diritto nella tecnica: tecnologie emergenti e nuove forme di regolazione*, in *federalismi.it*, XVI (2020), pp. 79-117.

Faini F., *Intelligenza artificiale e regolazione giuridica: il ruolo del diritto nel rapporto tra uomo e macchina*, in *federalismi.it*, 2, 2023, 1-29.

Finocchiaro G., *Diritto di Internet*, III ed., Zanichelli, Bologna, 2020.

Finocchiaro G., *Diritto dell'intelligenza artificiale*, Zanichelli, Bologna, 2024.

Fioriglio G., *La Società algoritmica fra opacità e spiegabilità: profili informatico-giuridici*, in «Ars Interpretandi», I (2021), pp. 53-67.

Floridi L., *La rivoluzione dell'informazione*, trad. it., Codice edizioni, Turin, 2012.

Frosini V., *Il diritto nella società tecnologica*, Giuffrè, Milan, 1981.

Frosini V., *L'orizzonte giuridico dell'Internet*, in «Il diritto dell'informazione e dell'informatica», II (2000), pp. 271-280.

Lagioia F., Sartor G., *L'intelligenza artificiale per i diritti dei cittadini: il progetto Claudette*, in «Ragion pratica», I (2020), pp.87-118.

Lessig L., *The Law of the Horse: What Cyberlaw Might Teach*, in *Harvard Law Review*, Vol. 113, II (1999), pp. 501-546.

Lessig L., *Code and Other Law of Cyberspace*, Basic Books, New York, 1999.

Llano-Alonso F.H., *L'etica dell'intelligenza artificiale nel quadro giuridico dell'Unione europea*, in Ragion Pratica, 2 (2021), pp. 327-348.

Maestri E., Lex informatica *e diritto. Pratiche sociali, sovranità e fonti nel cyberspazio*, in «Ars Interpretandi», I (2017), pp. 15-28.

Maestri E. (eds.), *Introduzione e note di sintesi al Regolamento (UE9 2024/1689 sull'intelligenza artificiale (testo ufficiale e integrale)*, Jovene, Napoli, 2024.

Micklitz H., Pollicino O., Reichman A., Simoncini A., Sartor G., De Gregorio G. (eds.), *Constitutional Challenges in the Algorithmic Society*, Cambridge University Press, Cambridge, 2021.

Moro P., *Algoritmi e pensiero giuridico. Antinomie e interazioni*, in «Rivista di diritto dei media», III (2019), pp.11-22.

Palmirani M., *Big Data e conoscenza*, in «Rivista di Filosofia del diritto», I (2020), pp, 73-92.

Pagallo U., *Algoritmi e conoscibilità*, in «Rivista di Filosofia del diritto», I (2020), pp. 93-106.

Pascuzzi G., *Il diritto dell'era digitale*, 6° ed., il Mulino, Bologna, 2025.

Punzi A., *Difettività e giustizia aumentata. L'esperienza giuridica e la sfida dell'umanesimo digitale*, in «Ars Interpretandi», I (2021), pp.113-128.

Rifkin J., *L'era dell'accesso. La rivoluzione della new economy*, trad. it., Mondadori, Milano, 2000.

Rodotà S., *Il diritto di avere diritti*, Laterza, Rome-Bari, 2012.

Sartor G., *L'informatica giuridica e le tecnologie dell'informazione. Corso di informatica giuridica*, II ed., Giappichelli, Torino, 2010.

Sartor G., *Internet e il diritto*, in C. Di Cocco, G. Sartor (eds.), *Temi di diritto dell'informatica*, IV ed., Giappichelli, Torino, 2020, 1 ff.

Sartor G., *L'intelligenza artificiale e il diritto*, Giappichelli, Torino, 2022.

Simoncini A., *Profili costituzionali della amministrazione algoritmica*, in «Rivista trimestrale di diritto pubblico», IV (2019), pp. 1149-1188.

Simoncini A., *Verso la regolamentazione della Intelligenza Artificiale. Dimensioni e governo*, in «BioLaw Journal», II (2021), pp. 411-417.

Simoncini A., Suweis S., *Il cambio di paradigma nell'intelligenza artificiale e il suo impatto sul diritto costituzionale*, in «Rivista di Filosofia del diritto», I (2019), pp. 87-106.

MARIA CRISTINA GAETA*

Risk assessment and management of Artificial Intelligence systems: the case study of brain computer inteface devices[§]

Abstract

This contribution analyses the process of assessing and managing damage caused by Artificial Intelligence, starting from the risk-based regulatory approach consistent with the provisions of the AI Act. Specifically, it examines the case study of the most sophisticated Brain-Computer Interface (BCI) devices, classified as high-risk Artificial Intelligence (AI). Despite the transformative potential of these technologies, the legal and ethical challenges posed by BCIs appear to be significant and relating to the violation of the rights of the individuals, as they are placed in a position of so-called technological vulnerability.

Il contributo analizza il processo di misurazione e gestione del danno da Intelligenza Artificiale, a partire dall'approccio normativo basato sul rischio, coerentemente con quanto disciplinato dall'AI Act. Nello specifico viene esaminato il caso dei più sofisticati dispositivi di Brain-Computer Interface (BCI), classificati come Intelligenza Artificiale (IA) ad alto rischio. Nonostante il potenziale trasformativo di tali tecnologie, appaiono significative le sfide giuridiche ed etiche poste dalle BCIs e relative alla violazione dei diritti degli individui, che è posto in una posizione di c.d. vulnerabilità tecnologica.

* Researcher (RTD-A) in Private Law, Coordinator of the Scientific Secretariat and External Relations of the Research Centre in European Private Law for Sustainable & Trustworthy Artificial Intelligence (ReCEPL4STAI), Università degli Studi Suor Orsola Benincasa (UNISOB). Legal expert for Technical Committee (CT) 324 '*Brain Computer* Interface' of the Italian Electrotechnical Committee (Comitato Elettrotecnico Italiano-CEI).

§ This contribution develops the presentation at the conference 'Cybersecurity e prodotti: sfide e responsabilità nelle società interconnesse', 29 November 2024, Italian Supreme Court.

1. The concept of 'risk' in the regulation of Artificial Intelligence

Artificial Intelligence, if properly regulated, is a clear strength for sustainable innovation. It is therefore necessary to promote responsible practices – from the design and implementation phase to the marketing and use of intelligent products – to ensure the protection of individuals with regard to technological development, placing AI technologies at the service of people and their well-being.

In this perspective, both from the point of view of prevention and repair of damage caused by AI to individuals, the classification of Artificial Intelligence based on risk *(risk-based approach)* and the system for *ma*naging such risks (risk management) are becoming increasingly central, as recently regulated in the AI Act (Reg. 2024/1689/EU)[1].

Given the rapid evolution of technologies, including AI, and their countless areas of application, it is not possible to define AI in a single type but, on the contrary, it must be appropriately classified. Among the different types of classification, the one based on risk is particularly relevant and topical because it is based on measuring the impact that these technologies can have on human beings and, in particular, on their rights. From an *ex ante* perspective, the correct classification of the AI system is the first step in verifying the need for a risk management system, with the provision of the relevant procedures and obligations. From an *ex post* perspective, on the other hand, risk management has obvious implications for liability, with reference to the identification of the responsible parties, the quantification of the damage and, more generally, the configurability of the liability itself, which is affected, among other things, by any product safety deficit, as also classified in the PLD proposal (text adopted on 12 March 2024).

From both perspectives, the now central role of the concept of measurement emerges, which is also of great importance in private law. It is necessary to measure the impact of AI technologies on human beings in order to correctly classify and manage risks, both through t the identification of adequate safeguards to prevent damage, such as those focused on legal subjectivity, contracting between private individuals and corporate social responsibility, as

[1] Regulation of the European Parliament and of the Council, 2024/1689/EU, 13.06.2024, laying down harmonised rules on Artificial Intelligence and amending Regulations No 300/2008/EC, No 167/2013(EU), No 168/2013/EU, No 2018/858/EU, No. 2018/1139/EU and No. 2019/2144/EU and Directives No. 2014/90/EU, No. 2016/797/EU and No. 2020/1828/EU (Artificial Intelligence Act – AI Act). Among the first to analyse the AI Act: G. Alpa, *Quale modello normativo europeo per l'Intelligenza Artificiale?,* in P.G. Monateri (ed.), *Scritti in memoria di Rodolfo*, UTET, Milan, 2024, pp. 1003-1026; R. Ruoppo, *Il regolamento UE in materia di Intelligenza Artificiale,* (c.d. AI Act): un quadro normativo uniforme per la tutela dei diritti fondamentali, in «Rass. dir. civ.», III (2024), pp. 995 ff.

well as through mechanisms for compensation and quantification of damage, with particular regard to strict liability, the provision of heavy financial penalties and so-called reputational penalties.

As for the type of regulation, given the complexity of the case in question, it can only be multi-level regulation (not only national, but also European and international), based on regulatory acts and techniques of varying effectiveness (hard law, soft law, standardisation, technoregulation).

2. Brain-computer interface devices as an example of high-risk AI

Neurotechnologies (NT) and, in particular, brain-computer interface (BCI) technologies, represent a rapidly evolving field with transformative potential for human health, well-being and interaction with the digital world, but also with significant ethical and legal implications. BCI devices include devices and procedures that monitor, analyse, predict or modulate the nervous system in order to understand, influence, restore or anticipate its structure, activity, function or intentions.

BCI technologies are a form of neural interface that enables direct communication between the brain and external devices, allowing, for example, the control of computers with thoughts or the restoration of motor and sensory functions. They are among the most promising technologies of the digital age, but they can entail multiple legal and ethical risks that must be analysed and, where possible, prevented. In particular, these innovations raise fundamental issues regarding self-determination, privacy, personal identity, freedom of thought and psychological manipulation.

In the relationship between humans and Artificial Intelligence technologies, such as the main BCI technologies, a new concept of 'vulnerability' emerges: technological vulnerability. This refers to the vulnerability of every human being with regard to technological development, regardless of any existing potential weakness (e.g. people with disabilities, minors, the elderly, individuals in positions of contractual power asymmetry or information asymmetry)[2], which implies the need for further investigation of the possible risks and specific protection.

[2] In this regard, L. Gatt, I.A. Caggiano, M.C. Gaeta, P. Grimaldi, A.A. Mollo, *Vulnerabilità tecnologica: una prospettiva di diritto privato*, in A. Patroni Griffi (ed.), *Vulnerabili. Questioni giuridiche, dilemmi bioetici,* Mimesis, Milan, 2025, 607-650, in which L. Gatt argues that 'Vulnerable subjects' refers to all persons acting in a technological environment, where they are more exposed to the risk of harm due to a 'vulnerability' that may exist for various reasons (minor age, asymmetry of bargaining power or information asymmetry). Its semantic scope includes all living beings – from humans to animals or plants (although this aspect is not considered in this document) – when they operate in contexts where other actors are more powerful for various reasons: social, economic or cultural, physical strength and, last

BCIs in general are characterised by strong public interest, and those in the consumer sector are becoming part of everyday life. However, this circumstance should not obscure the delicate nature of such devices, such as those in the medical field, which clearly demonstrate the urgent need for appropriate regulatory measures.

BCI devices are capable of establishing direct communication between the human brain and an external system (such as a computer or a robotic limb), opening up scenarios that seemed unthinkable just a few years ago. Think of a cursor on a screen moved simply by thinking, the possibility of adapting teaching methods to individual learning styles and monitoring students' attention. There are also neurotechnologies capable of improving the mobility of a paralysed person thanks to an exoskeleton guided by their brain impulses, as well as BCI technologies for the treatment of neurological disorders and to enable people with severe disabilities to express their will.

However, with technological progress, new legal and ethical challenges inevitably arise[3] that require adequate regulation, starting with dialogue between the humanities (in particular moral philosophy), social sciences (primarily law) and technical sciences (specifically engineering). BCIs are no exception. The ability to read and interpret brain activity raises questions about security and the resulting liability issues. Issues related to physical *safety* and Cybersecurity are also significant, such as attacks on the user and the BCI device, including via the internet *(brain hacking)*. With regard to privacy[4] , which is one of the main issues in the event of a breach of IT security, for example, the risk of unauthorised access to and processing of neural and

but not least, technological gap. The authors' study of these issues was developed within the Jean Monnet Chair '*European Protection Law of Individuals in Relation to New Technologies*' (PROTECH), established at Università Suor Orsola Benincasa, of which Lucilla Gatt is the holder. For further information, see also: L. Gatt, *Legal anthropocentrism between nature and technology: the new vulnerability of human beings*, in «European Journal of Privacy Law & Technologies», I (2022), pp. 15-26; L. Gatt, *The vulnerability of the human being in a technological environment: The need for protective regulation*, in L. Gatt (ed.), *Social networks and multimedia habitats,* Abstract Book, Suor Orsola Benincasa Univrsità Editrice, Naples, 2021, pp. 7-9. On technological vulnerability with specific reference to neurotechnologies, see also A.A. Mollo, *La vulnerabilità tecnologica. Neurorights ed esigenze di tutela: profili etici e giuridici,* in «European Journal of Privacy Law & Technologies», I (2021), pp. 199-210.

[3] R. Yuste, S. Goering, B. Arcas, *et al., Four ethical priorities for neurotechnologies and AI,* in «Nature», DLI (2017), pp. 159-163.

[4] For further information on the possible areas of application of neurotechnologies, including their use in the medical field by multinational companies and governments worldwide, with obvious consequences for user privacy, see N. Farahany, *Defending Our Brains*, Turin, 2023. I would also like to refer to M.C. Gaeta, *Brain Computer Interface technologies in medicine between regulation and future prospects,* in F. Trubiani (ed.), *Sistemi di intelligenza artificiale in medicina verso lo spazio europeo dei dati sanitari. Un dialogo multidisciplinare*, Giappichelli, Turin, 2025, pp. 255-273.

mental data[5], which could also be used for monitoring[6], psychological profiling and, in some cases, discrimination, is central. The security of BCI technologies also affects thought and decision-making, as BCI technologies can involve a loss or limitation of control over one's own mind, thus violating the user's right to self-determination through manipulative techniques. Therefore, there are also many issues related to civil liability, ranging from the difficulty of identifying the responsible party to proving and quantifying the damage suffered.

These are just some of the main risks that highlight the strong need for *'compliance by design'* that is both ethical and legal and centred on dialogue with engineers, those who are involved in the design and then implementation of such technologies. In fact, in order to address the challenges posed by new technologies, such as BCI, and to ensure safe neurotechnologies, it is essential to adopt a *compliance* approach from the design stage onwards, taking into account ethical principles, legal rules and how these principles and rules can be translated into programming language by engineers. This means integrating ethical and legal foundations into BCI systems from the design stage onwards, as well as ensuring compliance with technical standards. In other words, the protection of fundamental human rights must be an intrinsic element of the technology.

3. Institutional interventions to guarantee safe and reliable neurotechnologies

Given the complexity of the new legal issues arising from NT applications, including BCI, the need for *ad hoc* regulations for neurotechnologies is becoming increasingly clear. However, before binding regulations can be introduced, the first step is to analyse BCI applications in detail in order to understand how they work and the main risks involved. This is the strategy currently being adopted at the supranational and national level.

Specifically, at the international level, several institutions and organisations have recognised the need for a legal and ethical framework to guide the development and use of neurotechnologies.

[5] Advances in neuroscience and brain-computer interfaces allow access to and manipulation of neural information, making it necessary to establish clear guidelines to protect individuals' neural and mental privacy. See Garante per la protezione dei dati personali, *Privacy and neuro-rights. The person in the age of neuroscience,* Conference proceedings 28.01.2021; Press release from the Garante per la protezione dei dati personali, *Neuro-rights: Stanzione, Privacy Guarantor: defining a legal and ethical statute,* 13.05.2022.

[6] For a general study on the power of web surveillance, see S. Zuboff, *Il capitalismo della sorveglianza,* Luiss University Press, Rome, 2023.

In 2019, the Organisation for Economic *Co-operation* and Development (OECD) adopted the Recommendation on Responsible Innovation in Neurotechnology[7]. The aim is to guide governments and innovators in addressing the ethical, legal and social challenges raised by new neurotechnologies, while promoting innovation in the field of human health and well-being. The recommendation sets out specific principles[8] that are based on the potential of NT, but also on key issues relating to human identity, freedom of thought, autonomy, privacy and human well-being. Particular attention is paid to NT developed in the healthcare sector, which can pose particularly high risks to human beings, as in the case of the use of brain data, understood as data on the human brain, quantitative data on the structure, activity and function of the human brain, but also *neural data*, i.e. qualitative and quantitative data on the structure, activity and function of the nervous system, closely related to mental states and mental *data,* since all human cognitive and emotional activity is generated by the brain[9]. It is essential *to safeguard* such data, protecting it from unauthorised use, discrimination (in commercial, legal, employment or insurance contexts) and security breaches.

The United Nations Educational, Scientific and Cultural Organisation (UNESCO) plays a leading role in the international dialogue on the ethics of science and technology and is currently engaged in preparing the Recommendation on the Ethics of Neurotechnology[10]. The Recommendation also builds on the previous Recommendation on the Ethics of Artificial Intelligence, which anticipated ethical issues related to AI systems for neurotechnologies and *brain-computer* interfaces[11], now adopting a holistic, multicultural and multidisciplinary approach to systematically reflect on the impacts of NTs on

[7] OECD, *Recommendation of the Council on Responsible Innovation in Neurotechnology,* OECD/LEGAL/0457, 2019.

[8] The OECD Recommendation sets out nine principles: *1) Promote responsible innovation in neurotechnology to address health challenges; 2) Prioritise assessing safety in the development and use of neurotechnology; 3) Promote the inclusivity of neurotechnology for health; 4) Foster scientific collaboration in neurotechnology innovation across countries and disciplines; 5) Enable societal deliberation on neurotechnology; 6) Enable the capacity of oversight and advisory bodies to address novel issues in neurotechnology; 7) Safeguard personal brain data and other information gained through neurotechnology; 8) Promote cultures of stewardship and trust in neurotechnology across the public and private sector; 9) Anticipate and monitor potential unintended use and/or misuse of neurotechnology.*

[9] M. Ienca, J.J. Fins, R.J. Jox, *et al.*, *Towards a Governance Framework for Brain Data,* in «Neuroethics», Vol. 15, XX, 2022, pp.1-14.

[10] UNESCO, *Draft Recommendations on ethics of Neurotechnology,* SHS/BIO/AHEG-Neuro-2/2024/1, 2024.

[11] UNESCO, *Recommendation on the Ethics of Artificial Intelligence,* SHS/BIO/PI/2021/1, 2021. On the ethical aspects of NT, see also the UNESCO paper, *The Risks and Challenges of Neurotechnologies for Human Rights,* 2023.

human beings, societies and ecosystems. It promotes the development and responsible use of NT for the benefit of humanity, with a particular focus on damage prevention and equity of access, protecting the right to self-determination, freedom of thought and mental privacy. Significant importance is given to NT applications in relation to persons with physical and mental disabilities (including persons suffering from anxiety, depression, post-traumatic stress disorder, eating disorders, addictions, and victims of trauma and survivors of violence), providing them with support to contribute to the achievement of equal enjoyment of human rights. It is emphasised that Member States should encourage the development of NT applications in the medical field for the protection of health and, in general, the improvement of the lives of people with disabilities. The Recommendation also stresses the need for awareness of ethical challenges and proposes the development of an *Ethical Impact Assessment* (EIA) methodology and a research programme to monitor the ethical impacts of NT.

Furthermore, the *United Nations Human Rights Council* has published the 2024 report[12], which highlights that the progress of NTs – from those applied in the medical field to non-medical ones aimed at consumers – raises expectations and serious concerns, as they have a particularly significant impact on human rights and are capable of interfering with mental and cognitive functions. It promotes the responsible development and use of NT, considered '*a unique disruptive* technology', in a rapidly growing but not yet adequately regulated sector, in which *brain-computer interfaces* represent an important emerging category. The Recommendation also stresses the indispensability of a human rights impact assessment to be placed at the centre of the business model and strategies of neurotechnology companies. NT must be examined in the light of the fundamental principle of human dignity, assessing its ability to influence individuality, mental abilities and personality. With regard to freedom of thought, in particular, the United Nations Report highlights that current NT is capable of revealing real thoughts, acting as mind-reading devices. This raises concerns about neural data privacy, *hacking* and misuse for commercial purposes or predatory practices that also violate the right to personal integrity, as well as risks of NT being used as torture or in violation of the principle of due process. A human rights-based approach is therefore necessary to ensure fair and responsible access to NT, giving priority to those who need it, especially persons with disabilities, minors and the elderly and, more generally, persons with physical or mental health problems[13].

[12] UN Human Rights Council, *Impact, opportunities and challenges of neurotechnology with regard to the promotion and protection of all human rights,* Report, A/HRC/57/61, 2024.

[13] In addition to the international soft law instruments already mentioned, there is also the UNICEF working paper entitled '*Neurotechnology and Children*' from 2024, which focuses on the appli-

In line with international policies, the European Union has developed a regulatory approach to NTs and, in particular, BCIs, based on the principles of protecting fundamental rights and promoting responsible innovation, in accordance with what has already been adopted for AI in general. The aim is to balance the benefits of NT, especially in the medical field, with the potential risks in this sector, such as those to health, safety and individual autonomy. The European Parliament's 2024 study on mental privacy in neurotechnologies already focuses on social (e.g. the power of Big Tech), legal (in particular Cybersecurity and privacy) and ethical challenges (with specific reference to cognitive enhancement and psychological manipulation)[14], analysing the different types of NTs, starting with those in the medical field and then extending the analysis to those in the non-medical field. The European Union Council's Research Paper focused entirely on the promises and risks of BCI technologies as a new frontier in human-machine interaction[15]. BCIs are classified according to their degree of invasiveness (invasive, non-invasive and partially invasive)[16] and the types of interaction with the human brain (unidirectional and bidirectional)[17].

4. Risk assessment and management of BCI technologies: the role of standardisation

The analysis conducted so far on BCI applications clearly shows the need for adequate regulation to protect human beings in a broader perspective of sustainable development of new technologies.

The European protective approach can already be seen in the European Community era, and a clear example is the consumer protection legislation of

cations and implications of neurotechnologies for children and their rights, aiming to provide a preliminary contribution in this area.

[14] European Parliament, *The protection of mental privacy in the area of neuroscience,* study, PE 757.807, 2024.

[15] Council of the European Union, *From vision to reality. Promises and risks of Brain-Computer Interfaces,* Research Paper, 2024. For an initial analysis of the research paper, see A.A. Mollo, *The EU Council's September 2024 study on Brain-Computer Interface (BCI) technologies*, in «Persona e Mercato», IV, 2024, pp. 1522-1528.

[16] Invasive BCIs involve surgical procedures to implant electrodes directly into the brain or on its surface, offering high-resolution signals but with medical risks; non-invasive BCIs use external sensors (e.g. EEG), eliminating surgical risks but with weaker signals that are susceptible to interference; while partially invasive BCIs use minimally invasive surgery or existing medical technologies to insert electrodes into the brain. See Council of the European Union, From vision to reality. cit., 4.

[17] BCIs can be unidirectional – i.e. from the brain to the computer or vice versa, for stimulation – or bidirectional, allowing real-time, interactive communication. Currently, most are developed in clinical settings for therapeutic and rehabilitative treatments, but they can also be used for cognitive or sensory enhancement. See Council of the European Union, From vision to reality. cit., 4-5.

the 1990s, as well as the more recent regulations issued by the European Union: the Regulation on Artificial Intelligence (Reg. 2024/1689/EU-AI Act)[18] and the new regulation on producer responsibility (Reg. 2024/2853/EU-New PLD)[19], now expressly applicable to intelligent products such as BCIs.

In fact, neurotechnologies are also 'products' and, as such, may fall within the scope of existing consumer protection legislation when applied to relations between professionals and consumers. Over the years, changing needs for the protection of individuals in relation to the development and use of new technologies have required the introduction of new and stricter regulations which, starting from the concept of product 'safety', extend to that of 'risk' and is based on *risk impact assessment* and risk management of technological products, including those equipped with Artificial Intelligence (AI), precisely in order to reduce the risks to human beings[20].

More specifically, the safety requirement broadly encompasses both safety and (*Cyber*)security, both of which are already regulated. Safety also concerns individuals (users) who use neurotechnological applications, as well as other human beings and the surrounding environment in their interaction with individuals who use such neurotechnological applications. Cybersecurity, on the other hand, is related to the functioning of IT (Information Technology) systems and, in this context, Cybersecurity measures to protect smart products from *hacker* attacks and algorithmic *biases* are particularly important. Furthermore, adequate compliance with security also includes respect for privacy, as data must also be protected through appropriate technical and organisational me-

[18] For an in-depth analysis of the AI Act, see: A. Mantelero, G. Resta, G.M. Riccio (eds.), *Intelligenza artificiale commentario. Regolamento (UE) 2024/1689-AI ACT, Linee guida Commissione europea pratiche vietate, Disegno di legge sull'intelligenza artificiale,* Wolters Kluwer, Milan, 2025. For a more in-depth analysis of regulatory issues from an interdisciplinary perspective: L. Gatt, M.T. Loconte, M.E. Mazzarella, *L'ambito di applicazione, soggettivo e oggettivo*, in U. Ruffolo, A. Amidei, *Diritto dell'intelligenza artificiale*, Vol. 1, *Responsabilità. Contratto. Regolazione. Veicoli autonomi,* Luiss University Press, Rome, 2025, pp. 49-89. On the objectives pursued by the EU legislator with the AI Act: G. Finocchiaro, *I molteplici obiettivi perseguiti dal regolamento*, in F. Donati, G. Finocchiaro, F. Paolucci, O. Pollicino, *La disciplina dell'intelligenza artificiale*, Giuffrè, Milan, 2024, pp. 101-111; see also G. Cassano, E.M. Tripodi (eds.), *Il Regolamento Europeo sull'Intelligenza Artificiale*, Maggioli Editore, Rimini, 2024.

[19] Directive of the European Parliament and of the Council, 2024/2853/EU, 23.10.2024, on liability for defective products, repealing Council Directive 85/374/EEC.

[20] Under the old Directive 85/374/EEC (now repealed by Directive 2024/2853/EU), products (including intelligent products) must comply with existing safety requirements and standards. The European Commission has also developed an increasingly accurate reference framework that also includes the safety of consumer products and is supplemented by supervision by the competent public authorities. For the European Commission's legal framework on product safety, please refer to the official website: https://commission.europa.eu/business-economy-euro/doing-business-eu/eu-product-safety-and-labelling/product-safety_en

asures designed to ensure product safety in order to prevent intrusion, unauthorised use and manipulation of personal data.

Safety and the resulting liability in the event of damage due to non-compliance with product safety requirements have always been regulated at the European and national levels. Specifically, in the consumer sector, the aforementioned New PLD and, in Italy, the Consumer Code (in particular Articles 114 et seq. of the Consumer Code) and the Italian Civil Code (in particular Articles 2043 et seq. and Article 2059 of the Civil Code) are relevant.

Cybersecurity is similarly regulated at both EU and national level. This regulation includes the NIS2 Directive (Directive 2022/2555/EU, which replaced the previous NIS Directive 2016/114/EU, transposed in Italy by Legislative Decree No. 138/2024)[21], which aims to achieve a high common level of Cybersecurity throughout the European Union, ensuring the protection of smart products connected online. More recently, the EU Cyber Resilience Act (CRA – Reg. 2024/2847/EU), has been introduced, providing mandatory Cybersecurity standards for hardware and software products with digital elements (including AI products) and manufacturers must ensure products are secure by design, provide security updates, and report incidents.

This product safety legislation has been supplemented by a new package of EU Regulations based on the latest approach to risk assessment and management, ranging from the General Data Protection Regulation (Reg. 2016/679/EU-GDPR[22] , transposed in Italy by Legislative Decree No. 101/2018[23], which was incorporated into the amendments to the privacy code) to the AI Act. The regulatory framework on Artificial Intelligence liability seemed to be completed by the Directive on Artificial Intelligence Liability (AILD-COM(2022) 496 final)[24], which

[21] Directive of the European Parliament and of the Council, 2022/2555/EU, 14.12.2022, on measures for a high common level of Cybersecurity across the Union, amending Regulation No 910/2014/EU and Directive No 2018/1972/EU and repealing Directive No 2016/1148 (NIS 2 Directive).

[22] Regulation of the European Parliament and of the Council, 2016/679/EU, 27.04.2016, on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation).

[23] For an in-depth analysis of the GDPR: R. D'ORAZIO, V. CUFFARO, V. RICCIUTO (eds.), *I dati personali nel diritto europeo,* Giappichelli, Turin, 2019; F. Pizzetti, *Intelligenza artificiale, protezione dei dati personali e regolazione*, Giappichelli, Turin, 2018. One of the first applied research studies on European privacy reform, focusing on privacy consent in relation to processing by digital tools, including automated ones, was conducted by the Centre for Research in European Private Law (ReCEPL) at the Università degli Studi Suor Orsola Benincasa, and the results were published in L. Gatt, R. Montanari, I.A. Caggiano (eds.), *Privacy and Consent. A Legal and UX&HMI Approach*, Suor Orsola Benincasa Università Editrice, Naples, 2021. For an analysis of privacy in relation to the Internet of Things (IoT), please refer to M.C. Gaeta, *La protezione dei dati personali nell'Internet of Things: l'esempio dei veicoli autonomi*, in «Dir. inf. e informatica», I (2018), pp. 147-179.

[24] Proposal for a Directive of the European Parliament and of the Council on adapting the rules on non-contractual civil liability to Artificial Intelligence, COM/2022/496 final, (AI Liability Directive).

had the ambitious goal of harmonising the rules on non-contractual civil liability for AI damage across Europe. However, the proposal presented a number of critical issues, both in terms of the criteria for attributing liability and the harmonisation of national regulations and was withdrawn in February 2025.

Notably, with regard to the risk-based approach, in order to measure and manage the impact of AI on fundamental human rights, the AI Act provided for a risk management system for high-risk AI systems (Article 9 of the AI Act), as well as a fundamental rights impact assessment *(Fundamental Rights Impact Assessment*-FRIA) – pursuant to Article 27 – but only for high-risk AI systems expressly indicated[25]. This method of measurement and management based on the so-called *risk-based* approach, is increasingly at the centre of regulatory techniques and requires a multidisciplinary approach. Furthermore, this mechanism goes beyond the security requirements already in place at the European level, adapting regulation according to the level of risk of AI, as a result of measuring the impact of Artificial Intelligence technology on human rights, with a view to greater protection of the individual. A different level of risk would correspond to a more or less stringent *compliance by design* to be implemented before the commercialisation of BCI devices in order to verify their impact on those rights defined as 'neuro-rights'[26].

The provisions of the AI Act could effectively move in the direction of preventing certain risks to individuals, establishing a regulatory framework that must, however, be applied to technologies on a case-by-case basis. For example, BCI technologies may fall within the category of prohibited AI technologies (Article 5 of the AI Act) when they use subliminal techniques that act without a person's awareness or deliberately manipulative or deceptive techniques that materially distort human behaviour. In this regard, the Guidelines on Prohibited AI Systems (4.02.2025 C(2025) 844 final)[27] emphasi-

[25] For further information on the impact assessment on fundamental rights, see L. Gatt, I.A. Caggiano, M.C. Gaeta, E. Troisi, R. Savella, F. Pratesi, R. Trasarti, *FRIA implementation model according to the AI Act*, in «European Journal of Privacy Law & Technologies», II (2014), pp. 193-204. With specific reference to the impact on persons with disabilities, see A.A. Mollo, *Vulnerabilità e sostenibilità: primi spunti per uno studio dell'impatto sulle persone con disabilità e sulle future generazioni dei dispositivi [neuro]tecnologici,* in «European Journal of Privacy Law & Technologies», Special Issue (2023), pp. 28-40.

[26] The term 'neurorights' seems to have been used for the first time by M. Ienca, R. Andorno, *Towards New Human Rights in the Age of Neuroscience and Neurotechnology,* in «Life Sciences, Society and Policy», Vol. 13 (2017), pp. 1-27; then taken up again in M. Ienca, *On neurorights*, in «Frontiers in Human Neuroscience», Vol. 15 (2021), pp. 1-11 and in subsequent writings such as M. Ienca, J.J. Fins, R.J. Jox, *et al., Towards a Governance Framework for Brain Data*, cit.,pp. 1-14.

[27] European Commission Guidelines on prohibited Artificial Intelligence practices established by Regulation 2024/1689/EU (AI Act), 4.02.2025, C(2025) 844 final. More recently, see also A.A. Mollo, *La dichiarazione del Presidente del Garante Privacy italiano sui 'neurorights' del 30 mag-*

se the risks of neurotechnologies, making explicit reference to brain-computer interfaces, which increase the risk of sophisticated subliminal manipulation and the ability to influence human behaviour at the subconscious level, extending to advanced techniques such as *dream hacking* and *brain spyware*[28].

In fact, Joint Opinion No. 5/2021 of the European Data Protection Board and the European Data Protection Supervisor had already considered that the use of AI to track the emotions[29] of a natural person was totally inappropriate and should be prohibited, except in certain cases of use for health or research purposes, such as for example, with patients for whom emotion recognition could be relevant[30].

Law and ethics have the task of contributing to the regulation and definition of limits and responsibilities in the use of new technologies, en-

gio 2022: l'auspicio alla definizione di uno "statuto giuridico ed etico dei neurodiritti", in «Persona e mercato», II (2022), pp. 317-319; A.A. Mollo, *Neurorights. Una prospettiva di analisi interdisciplinare tra diritto e neuroscienze*, in S. Orlando, G. Capaldo (eds.), *Annuario OGID 2022,* Sapienza Università Editrice, Rome, 2022, pp. 191-215; A.A. Mollo, *Emessa in Cile il 9.8.2023 la prima sentenza al mondo sui neurodiritti (a proposito di 'Insight' un dispositivo neurotecnologico non terapeutico e non invasivo in commercio del tipo elettroencefalogramma mobile progettato per ottenere informazioni sull'attività cerebrale)*, in «Persona e Mercato», III (2023), pp. 603-605.

[28] The general regulation of smart products must then be supplemented by special regulations, which may be required for certain types of particularly risky products, such as BCIs and the aforementioned legislation currently being drafted. Furthermore, for some particularly complex areas, it is necessary to take into account the specific regulations, where applicable to the technology being regulated. For example, with regard to BCI devices, if applied in the medical-health field, European and national sector regulations must also be applied: Regulation of the European Parliament and of the Council 2017/745, 5.04.2017, on medical devices, amending Directive 2001/83/EC, Regulation No 178/2002/EC and Regulation No 1223/2009/EC and repealing Council Directives 90/385/EEC and 93/42/EEC (Medical Device Regulation-MDR); Regulation of the European Parliament and of the Council, 2025/327/EU, 11.02.2025, on the European Health Data Space and amending Directive 2011/24/EU and Regulation 2024/2847/EU; Law No. 59/2024, converting Decree Law No. 19/24, containing urgent provisions for the implementation of the PNRR, which amended Article 110 on privacy, on the legal basis for the processing of health data for scientific research in the medical, biomedical or epidemiological fields; Draft Law No. 1245/2024, Establishment of the National Authority for Artificial Intelligence and Neurotechnologies; Articles 7 and 8 of Law 23 September 2025, no. 132, *Disposizioni e deleghe al Governo in materia di intelligenza artificiale*, approving Draft Law No. 1146/2024, definitively approved on 17 September 2025 and published on 23 September 2025; Draft Law No. 1061/2024, Disposizioni per lo svolgimento di una sperimentazione in materia di utilizzo di dati sanitari.

[29] On the tracking of emotions of people with disabilities, see A.A. Mollo, *Emozioni e autodeterminazione in ambiente digitale: prime riflessioni sull'affect recognition e il neuromarketing relativi alle persone con disabilità,* in S. Orlando (ed.), *Profili giuridici del neuromarketing,* Annuario 2023-2024 OGID, Sapienza Università editrice, Rome, 2025, pp. 197-218.

[30] EDPB-EDPS, Joint Opinion 5/2021, 18 June 2021, *on the proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act).* In this regard, see S. Orlando, *Regole di immissione sul mercato e «pratiche di intelligenza artificiale» vietate nella proposta di Artificial Intelligence Act*, in «Persona e mercato», III (2022), pp. 346-367.

suring the protection of fundamental rights. The need for dialogue between law and ethics is particularly evident in certain areas of application, such as those characterised by the use of neurotechnologies[31].

In fact, although the legal system is not incomplete, it cannot regulate every single legal case, even if only probable, in the real world, in order to avoid excessive rigidity in the legal system, with overly stringent binding rules that hinder scientific, social and economic progress. At the same time, the obvious difficulty of keeping up with certain developments in real life, which are often faster than the legislative process, means that legislation comes into force already partially obsolete. This is the case, for example, with the regulation of new technologies, where the risk of regulatory obsolescence[32] is particularly high and adds to the more general issue of the rigidity of hard law regulation, which can also lead to multiple problems in relation to technological development. Consider the new technologies of AI and, even more so, those of BCI. In this context, the regulatory approach of the European legislator has mainly focused on the analysis of technologies through soft law acts (communications and studies by supranational institutions and organisations), and only more recently have the first regulatory proposals been published (such as Italian Draft Law No. 1245/2024), which, however, is a legislative proposal focused on identifying a competent authority for neurotechnologies rather than a regulatory act on them. Ethical principles can therefore be both a criterion for interpreting existing legislation, in accordance with current law, and a first non-binding regulatory intervention in a new legal area, possibly becoming binding s if incorporated into *hard law* regulations.

[31] A.A. Mollo, *Implicazioni etico-legali delle neurotecnologie: protezione dei dati, standard e regole di responsabilità,* in L. Gatt, M.C. Gaeta (eds.) *Building a New Strategy for European Protection Law in a technological environment,* Suor Orsola Benincasa Università Editrice, Naples, 2023, pp. 81–101; M.C. Gaeta, L. Aulino, E. Troisi, *The Possible Relationships Between Law and Ethics in the Context of Artificial Intelligence Regulation*, in «Humana.Mente Journal of Philosophical Studies», XLIV (2023), pp. 163-206.

[32] Regulatory obsolescence refers to the phenomenon whereby a law, regulation or rule becomes obsolete and no longer suited to the current reality due to the passage of time, technological progress, or social and cultural changes. See V. Cuffaro, *Obsolescenza e caducità delle leggi civili,* Roma Tre Press, Rome, 2019, which collects contributions on the effectiveness over time of civil laws which are intended to regulate private relationships of a patrimonial nature that, over a more or less long period of time, may not be adequate to the socio-economic or technical context of reference. The legal cases analysed concern the most classic ones of property law and contracts (e.g. industrial subcontracting, sale or network contracts), but the concept of regulatory obsolescence can be extended to all areas of application in which the case to be regulated and the context of reference evolve more rapidly than the regulatory framework.

With regard to the type of regulation, therefore, a multi-level regulatory approach is necessary, starting from a legal framework of binding law applicable at least at the European level (e.g. European Regulation), but supplemented by non-binding legislative acts characterised by the centrality of ethical principles, which complement and supplement the legal rules (e.g. guidelines, codes of conduct). Having identified the core of regulatory legislation, technical *standards* are of fundamental importance, especially if harmonised[33], as they play a central role in regulating more technical phenomena such as BCI, improving their safety (physical and IT, including privacy), interoperability with other devices and accessibility for users. In recent decades, as a result of technological development, the role of standardisation has become increasingly central and aims to identify technical specifications, including qualitative ones, with which manufacturers of AI systems (such as BCI technologies) can comply in order to ensure the harmonised safety and reliability of such systems, as well as their interoperability with other products or systems and their usability for the user.

Based on the regulatory framework described above, for the concrete compliance of BCI products, *compliance* mechanisms based on ethical principles and legal and technical rules can be identified, which operate from the design stage on a voluntary or legal basis and extend to the pre-market placing and periodic monitoring of such products. This leads to the development of the *compliance by design* mechanism, which includes the concrete verification of the impact of technologies on human rights and is key to ensuring the protection of individuals in relation to technological development, allowing the design and development of BCI technologies that serve humanity while minimising risks. This *compliance by design* can be implemented in practice through the use of *Legal&Ethical compliance by design tools*-designed to assess the impact of BCI devices on the most important human rights, with particular regard to physical and mental safety (including self-determination), pri-

[33] In the European regulatory framework, harmonised standards are technical documents drawn up by European standardisation bodies (CEN, CENELEC and ETSI) on behalf of the European Commission, which provide detailed technical specifications for products, processes or services in order to ensure compliance with European Union legislation. Conversely, non-harmonised standards are those that do not fall within this European framework and may be national or international in nature, but do not enjoy the presumption of conformity with EU legislation that derives from the publication of references to harmonised standards in the Official Journal of the European Union.

vacy and IT security-based on specific risk assessment parameters and measurement metrics[34] .

[34] A group of researchers from the Centre for Research in European Private Law (ReCEPL) at the Suor Orsola Benincasa University, together with the Ethics in the Wild Research Lab at the DH Centre of the University of Salento and the Dirpolis Institute at the Sant'Anna School of Advanced Studies in Pisa, has been studying *compliance by design* in technological devices, including BCI devices, for several years. In the field of BCI technologies, the research has been extended to the creation of a tool model for measuring the impact of AI technologies on fundamental human rights for Legal&Ethical compliance by design. For a detailed description of the studies conducted: L. Gatt, I.A. Caggiano, M.C. Gaeta, A.A. Mollo, *BCI Devices and their legal compliance: A prototype tool for its evaluation and measurement,* in «European Journal of Privacy Law & Technologies», I (2022), pp. 301-314; L. Gatt, I.A. Caggiano, E. Troisi, L. Aaulino, D.S. D'Aloia, L. Izzo, *BCI devices and their capacity to express human will having legal value: A model of risk-based classification,* in «European Journal of Privacy Law & Technologies», I (2022), pp. 398-300; L. Gatt, I.A. Caggiano, M.C. Gaeta, A.A. Mollo, *Neurorights in BCI Applications: a private law perspective,* in «European Journal of Law and Technology», Vol. 16, II (2025), pp. 1-14; M.C. Gaeta, F. Battaglia, *Legal & Ethical compliance by design for Brain Computer Interface devices,* in «CEI Magazine» IV (2025), accessible at the following link: https://ceimagazine.ceinorme.it/legal-ethical-compliance-by-design-per-i-dispositivi-di-brain-computer-interface/; F. Battaglia, G. Di Vetta, *Technology to unlock the mind: Citizen science and the sandbox approach for a new model of BCI governance,* in Proceedings of the IEEE MetroXRAINE 2022 conference, 2022, pp. 563-567.

BIBLIOGRAPHY

Alpa G., *Quale modello normativo europeo per l'Intelligenza Artificiale?*, in P.G. Monateri (ed.), *Scritti in memoria di Rodolfo*, UTET, Milan, 2024, pp. 1003-1026.

Battaglia F., Di Vetta G., *Technology to unlock the mind: Citizen science and the sandbox approach for a new model of BCI governance*, in Proceedings of the IEEE MetroXRAINE 2022 conference, 2022, pp. 563-567.

Cassano G., Tripodi E.M. (eds.), *Il Regolamento Europeo sull'Intelligenza Artificiale*, Maggioli Editore, Rimini, 2024.

Cuffaro V., *Obsolescenza e caducità delle leggi civili*, Roma Tre Press, Rome, 2019.

D'Orazio R., Cuffaro V., Ricciuto V. (eds.), *I dati personali nel diritto europeo*, Giappichelli, Turin, 2019.

Farahany N., *Defending Our Brains*, Turin, 2023.

Finocchiaro G., *I molteplici obiettivi perseguiti dal regolamento*, in Donati F., Finocchiaro G., Paolucci F., Pollicino O., *La disciplina dell'intelligenza artificiale*, Giuffrè, Milan, 2024, pp. 101-111.

Gaeta M.C., Battaglia F., *Legal & Ethical compliance by design for Brain Computer Interface devices*, in «CEI Magazine» IV (2025).

Gaeta M.C., Aulino L., Troisi E., *The Possible Relationships Between Law and Ethics in the Context of Artificial Intelligence Regulation*, in «Humana.Mente Journal of Philosophical Studies», XLIV (2023), pp. 163-206.

Gaeta M.C., *Brain Computer Interface technologies in medicine between regulation and future prospects*, in F. Trubiani (ed.), *Sistemi di intelligenza artificiale in medicina verso lo spazio europeo dei dati sanitari. Un dialogo multidisciplinare*, Giappichelli, Turin, 2025, pp. 255-273.

Gaeta M.C., *La protezione dei dati personali nell'Internet of Things: l'esempio dei veicoli autonomi*, in «Dir. inf. e informatica», I (2018), pp. 147-179.

Garante per la protezione dei dati personali, *Privacy and neuro-rights. The person in the age of neuroscience*, Conference proceedings 28.01.2021

Gatt L., Caggiano I.A., Gaeta M.C., Troisi E., Savella R., Pratesi F., Trasarti R., *FRIA implementation model according to the AI Act*, in «European Journal of Privacy Law & Technologies», II (2014), pp. 193-204.

Gatt L., Lo Conte M.T., Mazzarella M.E., *L'ambito di applicazione, soggettivo e oggettivo*, in Ruffolo U., Amidei A., *Diritto dell'intelligenza artificiale*, Vol. 1, *Responsabilità. Contratto. Regolazione. Veicoli autonomi*, Luiss University Press, Rome, 2025, pp. 49-89.

Gatt L., Caggiano I.A., Troisi E., Aulino L., D'Aloia D.S., Izzo L., *BCI devices and their capacity to express human will having legal value: A model of risk-based classification*, in «European Journal of Privacy Law & Technologies», I (2022), pp. 398-300.

Gatt L., Caggiano I.A., Gaeta M.C., Mollo A.A., *BCI Devices and their legal compliance: A prototype tool for its evaluation and measurement*, in

«European Journal of Privacy Law & Technologies», I (2022), pp. 301-314.

Gatt L., Caggiano I.A., Gaeta M.C., Mollo A.A., *Neurorights in BCI Applications: a private law perspective*, in «European Journal of Law and Technology», Vol. 16, II (2025), pp. 1-14.

Gatt L., Caggiano I.A., Gaeta M.C., Grimaldi P., Mollo A.A., *Vulnerabilità tecnologica: una propsettiva di diritto privato*, in Patroni Griffi A. (ed.), *Vulnerabili. Questioni giuridiche, dilemmi bioetici*, Mimesis, Milan, 2025, 607-650.

Gatt L., Montanari R., Caggiano I.A. (eds.), *Privacy and Consent. A Legal and UX&HMI Approach*, Suor Orsola Benincasa Università Editrice, Naples, 2021.

Gatt L., *Legal anthropocentrism between nature and technology: the new vulnerability of human beings*, in «European Journal of Privacy Law & Technologies», I (2022), pp. 15-26.

Gatt L., *The vulnerability of the human being in a technological environment: The need for protective regulation,* in L. Gatt (ed.), *Social networks and multimedia habitats*, Abstract Book, Suor Orsola Benincasa Univrsità Editrice, Naples, 2021, pp. 7-9.

Ienca M., Andorno R., *Towards New Human Rights in the Age of Neuroscience and Neurotechnology*, in «Life Sciences, Society and Policy», Vol. 13 (2017), pp. 1-27.

Ienca M., Fins J.J., Jox R.J.-et al., *Towards a Governance Framework for Brain Data*, in «Neuroethics», Vol. 15, XX (2022), p. 1-14.

Ienca M., *On neurorights*, in «Frontiers in Human Neuroscience», Vol. 15 (2021), pp. 1-11.

Mantelero A., Resta G., Riccio G.M. (eds.), *Intelligenza artificiale commentario. Regolamento (UE) 2024/1689-AI ACT, Linee guida Commissione europea pratiche vietate, Disegno di legge sull'intelligenza artificiale*, Wolters Kluwer, Milan, 2025.

Mollo A.A., *Emessa in Cile il 9.8.2023 la prima sentenza al mondo sui neurodiritti (a proposito di 'Insight' un dispositivo neurotecnologico non terapeutico e non invasivo in commercio del tipo elettroencefalogramma mobile progettato per ottenere informazioni sull'attività cerebrale),* in «Persona e Mercato», III (2023), pp. 603-605.

Mollo A.A., *Emozioni e autodeterminazione in ambiente digitale: prime riflessioni sull'affect recognition e il neuromarketing relativi alle persone con disabilità*, in S. Orlando (ed.), *Profili giuridici del neuromarketing*, Annuario 2023-2024 OGID, Sapienza Università editrice, Rome, 2025, pp. 197-218.

Mollo A.A., *Implicazioni etico-legali delle neurotecnologie: protezione dei dati, standard e regole di responsabilità*, in L. Gatt, M.C. Gaeta (eds.) *Building a New Strategy for European Protection Law in a technological environment*, Suor Orsola Benincasa Università Editrice, Naples, 2023, pp. 81-101.

Mollo A.A., *La dichiarazione del Presidente del Garante Privacy italiano sui 'neurorights' del 30 maggio 2022: l'auspicio alla definizione di uno "statuto giuridico ed etico dei neurodiritti"*, in «Persona e mercato», II (2022), pp. 317-319.

Mollo A.A., *Neurorights. Una prospettiva di analisi interdisciplinare tra diritto e neuroscienze*, in S. Orlando S., G. Capaldo (eds.), *Annuario OGID 2022*, Sapienza Università Editrice, Rome, 2022, pp. 191-215.

Mollo A.A., *La vulnerabilita tecnologica. Neurorights ed esigenze di tutela: profili etici e* giuridici, in «European Journal of Privacy Law & Technologies», I (2021), pp. 199-210.

Mollo A.A., *Vulnerabilità e sostenibilità: primi spunti per uno studio dell'impatto sulle persone con disabilità e sulle future generazioni dei dispositivi [neuro]tecnologici*, in «European Journal of Privacy Law & Technologies», Special Issue (2023), pp. 28-40.

Orlando S., *Regole di immissione sul mercato e «pratiche di intelligenza artificiale» vietate nella proposta di Artificial Intelligence Act*, in «Persona e mercato», III (2022), pp. 346-367.

Pizzetti F., *Intelligenza artificiale, protezione dei dati personali e regolazione*, Giappichelli, Turin, 2018.

Ruoppo R., *Il regolamento UE in materia di Intelligenza Artificiale, (c.d. AI Act): un quadro normativo uniforme per la tutela dei diritti fondamentali*, in «Rass. dir. civ.», III (2024), pp. 995 ff.

Yuste R., Goering S., Arcas B.-et al., *Four ethical priorities for neurotechnologies and AI*, in «Nature», DLI (2017), pp. 159-163.

Zuboff S., *Il capitalismo della sorveglianza*, Luiss University Press, Rome, 2023.

ANDREA GENTILE*

Cybersecurity and Criminal Law: Aspects of Corporate§

Abstract

The rapid development of Artificial Intelligence and the increasing relevance of Cybersecurity are challenging traditional categories of criminal law and corporate liability. This paper analyses how criminal law must adapt to technologically driven risks while preserving its fundamental guarantee-based principles. The first part examines the implications of AI for criminal responsibility, focusing on attribution of conduct, subjective liability, and the tension between algorithmic autonomy and the principle of personal responsibility, as well as issues of transparency and the right to defense. The second part addresses Cybersecurity and corporate criminal liability in light of recent Italian and European reforms, particularly Law No. 90/2024 and Legislative Decree No. 138/2024 implementing the NIS2 Directive. Special attention is devoted to the expansion of corporate liability under Legislative Decree No. 231/2001 and the consequent need to revise organizational and compliance models. The paper concludes by advocating an integrated, preventive, and risk-based regulatory approach capable of reconciling technological innovation with effective protection of fundamental rights.

Il rapido sviluppo dell'Intelligenza Artificiale e la crescente rilevanza della Cybersicurezza mettono in discussione le categorie tradizionali del diritto penale e della responsabilità degli enti. Il contributo analizza le modalità attraverso cui il diritto penale è chiamato ad adattarsi ai rischi tecnologici, nel rispetto dei principi fondamentali di garanzia. La prima parte esamina le implicazioni dell'IA in tema di responsabilità penale, con particolare riferimento

* Professor of Criminal Law at the San Raffaele Telematic University of Rome.

§ This contribution develops the presentation at the conference 'Cybersecurity e prodotti: sfide e responsabilità nelle società interconnesse', 29 November 2024, Italian Supreme Court.

all'attribuzione della condotta, all'imputazione soggettiva e alla tensione tra autonomia algoritmica e principio di responsabilità personale, nonché ai profili di trasparenza e tutela del diritto di difesa. La seconda parte affronta il tema della Cybersicurezza e della responsabilità da reato degli enti alla luce delle recenti riforme italiane ed europee, in particolare la l. n. 90/2024 e il d.lgs. n. 138/2024 di attuazione della direttiva NIS2. L'attenzione si concentra sull'ampliamento della responsabilità ex d.lgs. n. 231/2001 e sulla conseguente esigenza di aggiornare i modelli organizzativi e di compliance. Il contributo si conclude sostenendo la necessità di un approccio integrato, preventivo e basato sul rischio, capace di conciliare innovazione tecnologica e tutela effettiva dei diritti fondamentali.

1. Introduction

Technological and digital evolution imposes a profound transformation in social and economic dynamics, which must lead scholars and legal practitioners to question how criminal law should adapt to these developments.

Criminal law, traditionally rooted in classical conceptual frameworks and particularly resistant to technological innovations, today faces phenomena that challenge its foundations: Artificial Intelligence (AI), with its relatively autonomous capacity to make decisions and impact reality, and Cybersecurity, which is linked to the growing exposure of businesses to cyber threats. The analysis first focuses on the relationship between AI and criminal law, and then explores the connection between Cybersecurity and the protection of enterprises, in light of the latest Italian and European regulations, in particular L. 90/2024 (so-called DDL Cybersicurezza) and D.lgs. 138/2024 (implementation of the NIS2 directive). The focus is primarily on criminal law and organizational models.

The purpose of this paper is to highlight how the law is called upon to undergo a constant process of adaptation in order to respond to digital risks while at the same time safeguarding fundamental rights.

It is clear that law — as a science called upon to resolve, in a broad sense, conflicts within the legal order in the face of emerging issues — cannot avoid involving some of its own branches, and in particular criminal law, which, in this sense, is specifically called upon by the relentless and pressing innovations of Artificial Intelligence to rethink classical frameworks and develop new organizational and behavioral models[1].

[1] See F. Basile, B. Fragasso, *Intelligenza artificiale e diritto penale, prove tecniche di convivenza*, in «Diritto di difesa», (2023), pp. 752-774.

The topic of the relationship between Artificial Intelligence and criminal law has been described by some observers as a true epochal challenge: AI is not merely a technological tool, but a set of algorithms and processes capable of learning, evolving, and autonomously impacting reality. Such a characteristic raises numerous problems of criminal liability.

In this regard, it should be noted that while on the one hand criminal law cannot renounce its role of safeguarding particularly significant legal interests, on the other, the process of rethinking its classical frameworks must take place with respect for the consolidated principles of guarantees recognized by our legal system.

A central point concerns the principle of personal responsibility, which in the distinctive sense of modern criminal law does not merely entail the prohibition of liability for the acts of others, but implies that the criminal act must be attributable as "one's own" to the acting subject: it follows that each unlawful act must correspond to a human author, conscious and endowed with the capacity to understand and to will. However, the decision-making autonomy of AI systems risks diluting this connection, generating gray areas in which it becomes difficult to establish who is responsible for harmful conduct. This raises the question of the attribution of the offense: is it the programmer who wrote the algorithm, the user who employs it, the company that implemented it, or none of these subjects in the traditional sense?

Another issue concerns the categories of intent and negligence: conduct generated by AI may be unpredictable for its creator, making it difficult to distinguish between human error and algorithmic autonomy[2]. This forces a reconsideration of the criteria for subjective imputation, also taking into account the predictability of the risk and the level of preventive measures adopted by those using such tools.

Finally, of particular importance is the issue of transparency and the protection of the right to defense. Where a judicial decision or procedural act is based on AI systems, it becomes essential to ensure the traceability and comprehensibility of the algorithms employed[3]. The risk is that criminal proceedings may rest on opaque and unverifiable processes, with serious consequences in terms of substantive justice[4].

[2] From this perspective, in order to properly delimit the scope of criminal liability, it is necessary to determine the so-called permitted risk. With regard to this matter see F. Basile, B. Fragasso, *Intelligenza artificiale*, cit., *passim*.

[3] Legal scholarship, with regard to the possible use of A.I. systems in judicial proceedings, highlights the need for what is defined as meaningful human control. *Ex multiis* G. Ubertis, *Intelligenza artificiale, giustizia penale, controllo umano significativo,* in «Diritto penale contemporaneo – Rivista trimestrale», IV (2020).

[4] A. Balsamo, *L'impatto dell'intelligenza artificiale nel settore della giustizia*, in «Sistema penale», (2024), p. 7 among the main disadvantages of using algorithms, it points out the black box problem, namely the opacity of the algorithm's functioning.

2. Cybersecurity and corporate criminal liability

Alongside the debate on Artificial Intelligence, Cybersecurity today represents one of the most delicate sectors of criminal law and corporate governance. Companies operate in a globalized and interconnected context, exposed to constant threats ranging from targeted cyberattacks on critical infrastructures to ransomware, and even sophisticated industrial espionage activities.

The Italian legislator, following European regulations, has introduced important measures aimed at strengthening the country's Cybersecurity. In particular, l. 28 giugno 2024, n. 90 (DDL Cybersicurezza), precisely in order to ensure greater protection through the deterrent function of criminal sanctions, has carried out a true organic reform of computer crimes: first of all, it increased statutory penalties for cases of unlawful access, illegal interceptions, and data damage. The reform abolished art. 615-quinquies c.p., followed by its substantial relocation into the new art. 634-quater.1, which was toughened and enriched with new aggravating circumstances, now classified among the crimes of data and computer system damage[5].

The legislator's intervention finds its rationale in the need to align the penal response with the seriousness and growing frequency of cyberattacks[6].

The Bill also introduced a new form of cyber extortion (art. 629, co. 3 c.p.), modeled on ransomware cases. This new offense must be welcomed, given that prior to its introduction, the conduct described, though punishable, lacked a unified criminal relevance, with a consequent risk of fragmentation and dispersion of law enforcement activities[7]. Part of legal scholarship[8] however, criticized the absence, alongside art. 629, co. 3 c.p., of concrete measures for broader investigative coordination. This gap appears even more significant considering that such crimes are often committed within complex international organizations.

[5] S. Pietropaoli, *Un'occasione (forse) mancata. Considerazioni sulla revisione dei reati informatici proposta con il DDL Cybersicurezza*, in «Rivista italiana di informatica e diritto», I (2024), pp.47-53 it recognizes that the relocation is supported by a "thematic" justification.

[6] A. Renzi, *Proteggere le nuove frontiere del paese, un'analisi della nuova legge sulla cybersicurezza nazionale italiana*, in «Azienditalia», (2024), pp. 1110-1117.

[7] G.Fiorinelli, *Il ransomware nel DDL Cybersicurezza: dalla fattispecie di estorsione "informatica" al coordinamento tra indagini e incident response*, in «Rivista italiana di informatica e diritto», I (2024), pp. 35-39. which, while on the one hand praises the legislator's intention since it deems it at least difficult to trace the conduct sanctioned by the new co. 3 of art. 629 back to the model of 'ordinary' extortion, on the other hand would have preferred the recovery of the model of an offense committed for the purpose of extortion.

[8] G. Corasaniti, *Strategie di contrasto al Ransomware e nuove frontiere della criminalità informatica,* in «Diritto dell'informazione e dell'informatica», II (2025), pp.19-46.

Finally, the reform introduced an extension of specific intent, the introduction of new aggravating circumstances, and the exclusion of mitigating factors for conduct committed using IT tools to obtain undue advantages at the expense of others, or to unlawfully access computer systems and/or intercept or disrupt computer and telematic communications[9].

From a procedural standpoint, art. 17 of the D.L. amended the Code of Criminal Procedure, whereby through the innovation of art. 51, co. 3-quinquies c.p.p., jurisdiction over investigations into certain computer crimes was assigned to the district prosecutor's office. Furthermore, through the modification of artt. 406, co. 5-bis and 407, co. 2 lett. a) c.p.p., the exceptional regime for extending preliminary investigations and the maximum two-year duration of preliminary investigations was extended to the crimes in question[10]. The rules on wiretapping applicable to organized crime were also made applicable to certain computer crimes.

These changes have been welcomed by scholars because they successfully delineate a model that recognizes the complexity of cyber investigations and promotes structured collaboration among different institutions, overcoming previous investigative fragmentation[11].

The D.lgs. 4 settembre 2024, n. 138 implemented the European NIS2 directive, considerably expanding the range of entities subject to Cybersecurity obligations. In addition to companies operating in critical sectors, the legislation also involves many business entities previously excluded. Among the obligations introduced are registration with the National Cybersecurity Agency (ACN), the establishment of risk management systems, timely notification of incidents, and mandatory internal training. Penalties for non-compliance can reach up to 10 million euros or 2% of global annual turnover.

3. D.L. Cybersecurity and the impact on corporate crime liability

One of the most innovative aspects of l. 90\2024 concerns its impact on the system of corporate criminal liability under D.lgs. 231/2001. The legislator extended this type of liability in the field of cybercrime, driven by the need to overcome difficulties related to individual attribution, which in this context proves hard to ascertain due to the delocalization of conduct.

[9] S. Liistro, *Legge 90\2024 sulla Cybersicurezza e compliance integrata: gli impatti su Modello 231 e privacy,* in «Norme e Tributi plus – il Sole 24 ore», (2024).

[10] G. Fiorinelli, *Il ransomware*, cit., 38 it points out the failure to refer to the art. 629, co. 3, of the criminal code in the provisions of the D.L. Cybersecurity on Cybersecurity that amend the c.p.p.

[11] G. Vaccari, *Cybersecurity nell'attività d'impresa, profili penalistici,* in www.altalex.it.

This makes it complex to identify the locus commissi delicti, preventing the certain allocation of territorial jurisdiction[12]. Secondly, empirical evidence shows that in cybercrime the harmful event is caused by a series of technical and automated events, in which human conduct occurs upstream, making it difficult to determine the causal link between the human act and the harmful event.

The law significantly amended art. 24-bis, providing in co. 1 for a general increase in monetary penalties, which now range from a minimum of 500 to a maximum of 700 units, replacing the previous range of 100 to 200.

The legislator also introduced, in co. 1-bis, the new offense of cyber extortion, punishable with a fine ranging from 300 to 800 units, accompanied by interdictive sanctions of at least two years.

The inclusion of this newly created offense among the predicate crimes of corporate administrative liability appears inconsistent with criminological data, given that it is highly unlikely that such attacks would be carried out by economic entities (the main addressees of 231 liability), as these activities are typically attributable to criminal organizations[13].

The legislator also replaced the repealed Art. 615-quinquies by introducing, through art. 24-bis, the "new" Art. 635-quater.1. These provisions significantly extend criminal risk for companies, even beyond traditionally technological sectors.

It should not be overlooked that several cybercrimes simultaneously involve personal data breaches, with the consequent obligation to notify the Data Protection Authority within 72 hours under Art. 33 of the GDPR. This area features profound interconnections between D.lgs. 231, compliance and privacy regulations, requiring companies to adopt an integrated compliance approach.

Case law reported by scholars[14] shows that such risks may also arise in non-tech contexts: unauthorized access to databases by former employees, tampering with video surveillance systems to conceal liability in workplace safety matters, etc. In all these cases, the risk for the entity is not merely theoretical but concrete and significant.

In light of this, companies are called upon to update their Organizational Models under D.lgs. 231\2001 with specific measures such as:

— Strict segregation of logical access;

— Continuous monitoring systems;

— Training and awareness programs for employees at all levels, including dedicated training sessions;

[12] *Ivi.*

[13] G. Fiorinelli, cit., p. 38.

[14] S. Liistro, cit.

— Integration between the 231 Model and the Privacy Model;

— Adoption of a risk-based approach capable of preventing cyber threats and fostering a corporate culture of digital security.

4. Critical considerations and future perspectives

Criminal law, challenged by the development of Artificial Intelligence and the growing relevance of Cybersecurity, must face new and complex issues. AI requires a rethinking of some of its categories and greater attention to algorithmic transparency. Cybersecurity, on the other hand, demands strong accountability of companies and their top management, imposing a profound revision of 231 organizational models. The shared intent is to harmonize innovation and safeguards, preventive and repressive policies, so that technological progress does not translate into new legal gaps, but instead becomes an opportunity to strengthen the protection of rights and collective trust.

Traditional criminal law, to fulfill its deterrent function, must be embedded in a legal system characterized by the presence of civil and administrative risk prevention and management tools. A legal model based primarily on repressive-deterrent protection would be less effective than a system based on prevention[15], which would allow criminal penalties to repress at the end of the chain only those particularly insidious behaviors that preventive measures failed to intercept upstream[16].

The current regulatory phase demonstrates an important effort by the Italian and supranational legislator to govern complex and constantly evolving phenomena. However, some critical issues remain. Regarding AI, the main difficulty lies in building legal categories capable of capturing diffuse and multi-level responsibilities without sacrificing the fundamental principles of criminal law. On the Cybersecurity front, the greatest problem lies in the global dimension of cyber threats: servers located in foreign jurisdictions, transnational conduct, and complex causal chains often make it difficult to establish jurisdiction and attribution.

For these reasons, the future of criminal law and corporate compliance will inevitably be characterized by growing international cooperation, constant interdisciplinary dialogue among law, technology, and economics, and a cultural approach that conceives security not as a cost but as a strategic investment and legal duty.

[15] The regulatory model adopted by our legislator through the D.lgs. 138/2024.

[16] Along these lines P.G. Chiara, *DDL Cybersicurezza: tra l'inasprimento della risposta penale del legislatore nazionale e il modello preventivo-amministrativo della direttiva NIS2*, in «Rivista italiana di informatica e diritto», I (2024), pp. 31-34.

BIBLIOGRAPHY

Balsamo A., *L'impatto dell'intelligenza artificiale nel settore della giustizia*, in «Sistema penale», (2024).

Basile F., Fragasso B., *Intelligenza artificiale e diritto penale, prove tecniche di convivenza*, in «Diritto di difesa», (2023), pp. 752-774.

Chiara P.G., *DDL Cybersicurezza: tra l'inasprimento della risposta penale del legislatore nazionale e il modello preventivo-amministrativo della direttiva NIS2*, in «Rivista italiana di informatica e diritto», I (2024), pp. 31-34.

Corasaniti G., *Strategie di contrasto al Ransomware e nuove frontiere della criminalità informatica*, in «Diritto dell'informazione e dell'informatica», II (2025), pp.19-46.

Fiorinelli G., *Il ransomware nel DDL Cybersicurezza: dalla fattispecie di estorsione "informatica" al coordinamento tra indagini e incident response*, in «Rivista italiana di informatica e diritto», I (2024), pp. 35-39

Liistro S., *Legge 90\2024 sulla Cybersicurezza e compliance integrata: gli impatti su Modello 231 e privacy*, in «Norme e Tributi plus – il Sole 24 ore», (2024).

Pietropaoli S., *Un'occasione (forse) mancata. Considerazioni sulla revisione dei reati informatici proposta con il DDL Cybersicurezza,* in «Rivista italiana di informatica e diritto», I (2024), pp.47-53.

Renzi A., *Proteggere le nuove frontiere del paese, un'analisi della nuova legge sulla cybersicurezza nazionale italiana*, in «Azienditalia», (2024), pp. 1110-1117.

Vaccari G., *Cybersecurity nell'attività d'impresa, profili penalistici*, in www.altalex.it.

PATRIZIA GIUSTI*

Relationships and interconnections between Cybersecurity and corporate liability: the "derived" obligation of crime risk prevention§

Abstract

The entry into force of Law No. 90/2024 on Cybersecurity raises significant questions regarding its interaction with existing corporate compliance frameworks, particularly Legislative Decree No. 231/2001 on the administrative liability of entities. This paper examines the structural connections between the new Cybersecurity regime and corporate criminal liability, focusing on the expansion of predicate cyber offenses, the strengthening of sanctions, and the resulting compliance obligations for organizations. Special attention is devoted to the role of risk assessment and organizational models as preventive tools, highlighting how Cybersecurity and corporate liability increasingly rely on a shared methodological approach. The analysis argues that, although formally non-mandatory, the adoption and updating of organizational, management, and control models under Decree No. 231/2001 have become de facto compulsory for entities subject to the Cybersecurity legislation. The paper concludes by stressing the need for greater coordination among Cybersecurity, corporate liability, and broader compliance regulations.

L'entrata in vigore della l. n. 90/2024 in materia di Cybersicurezza solleva rilevanti interrogativi circa la sua interazione con i preesistenti sistemi di compliance aziendale, in particolare con il d.lgs. n. 231/2001 sulla responsabilità da reato degli enti. Il contributo analizza i collegamenti strutturali tra la nuova disciplina della Cybersicurezza e la responsabilità dell'ente, soffermandosi sull'ampliamento dei reati presupposto, sull'inasprimento del siste-

* Lawyer, Criminologist, Internet Security Consultant, GDPR and Data Protection Officer, and Member of Supervisory Bodies (Organismo di Vigilanza).

§ This contribution develops the presentation at the conference 'Cybersecurity e prodotti: sfide e responsabilità nelle società interconnesse', 29 November 2024, Italian Supreme Court.

ma sanzionatorio e sulle conseguenti ricadute organizzative per le imprese. Particolare attenzione è dedicata al ruolo della valutazione del rischio e dei modelli organizzativi quali strumenti di prevenzione, evidenziando come Cybersicurezza e responsabilità d'impresa condividano un comune approccio metodologico. L'analisi sostiene che l'adozione e l'aggiornamento del Modello 231, pur non formalmente obbligatori, assumano oggi un carattere sostanzialmente necessario per i soggetti destinatari della normativa sulla Cybersicurezza.

1. Introduction

With the entry into force of the DDL Cybersecurity, we wonder if and how the new discipline will impact other already existing regulations dictated for corporate compliance.

The subject of Cyber security requires an integrated approach.

This paper will focus on the connections with D. Lgs. 231/2001, which, as is well known, governs the administrative liability of entities, companies, and associations, including those without legal personality.

A law that has changed over time and which today must live with situations that did not exist in 2001.

Finally, the paper will explore how the two provisions — so different in their rationale but united by a shared goal — will manage to avoid being yet another paper-based obligation and instead become useful strategic tools for the success of organizations and businesses.

2. New Cybercrime law (DDL)

The Annual Report Parliament presented by the National Cybersecurity Agency (ACN) highlighted that in 2022 alone, over a thousand Cybersecurity incidents were handled; a number underestimated, considering the huge number of cyber incidents, for reasons of image and liability[1].

On June 19, 2024, the Senate approved Bill No. 1717, which was adopted with Law No. 90 of June 28, 2024, containing Provisions for Strengthening National Cybersecurity and Cybercrime (the so-called Cybersecurity Law), which came into force on July 17[2].

[1] Can.gov.it/portale/relazione-annuale.

[2] https://www.normattiva.it/uri-res/N2Ls?urn:nir:stato:legge:2024-06-28.

The law, composed of just two chapters and a total of 24 articles, is dedicated — in its first part — its dedicated to strengthening national Cybersecurity and cybercrime.

The second part is devoted to sanctions of cybercrimes, with significant amendments to the criminal code, the criminal procedure code, and the regulations concerning the liability of entities. It also includes the strengthening of the functions of the National Cybersecurity Agency (ACN) and its coordination with the judiciary in cases of cyberattacks, with the introduction of specific procedures aimed at enabling quicker intervention by the Agency, for the purpose of preventing attacks and their consequences and restoring the functionality of information systems[3].

3. The new features of the law

The DDL has four main areas of intervention:[4]

– *Obligation to notify the Agency for Cybersecurity* incidents, which applies to central public administrations, regions, provinces, metropolitan cities, regional capital municipalities, municipalities with populations over 100,000 inhabitants, local health authorities, public transport companies, in-house companies providing IT services, transport, waste collection, disposal, or wastewater treatment. Failure to comply with this obligation may result in penalties ranging from €25,000 to €125,000, and the public administration employee found to have violated this provision may face disciplinary and administrative-accounting responsibility proceedings.

– *Important changes to cybercrime offenses*, which significantly impact various aspects, including:

– *Introduction of an aggravating circumstance for fraud* when committed remotely, using IT or telematics tools that hinder identification, with the extension of mandatory confiscation under Article 240, paragraph 2 of the Penal Code of the goods and IT or telematics instruments.

– *Extension of the aggravating circumstance for unauthorized access to an IT or telematics system* (Art. 615-ter c.p.) when the act results in the theft, even through reproduction or transmission, or inaccessibility to the owner.

– *Expansion of the offense of unauthorized possession, dissemination, and installation of equipment, codes, and other means to access IT or telema-*

[3] https://www.governo.it/articolo/comunicato-stampa-del-consiglio-dei-ministri-n-66/24831.

[4] https://www.4clegal.com/opinioni/rafforzamento-cybersecurity-nazionale-legge-entrera-vigore-17 luglio#:~:text=Abstract,principali%20novit%C3%A0%20e%20alcune%20considerazioni.&text=Il%2017%20luglio%202024%20entrer%C3%A0%20in%20vigore%20la%20L.

tics systems (Art. 615-quater c.p.) to cases where the act is committed with the intent to gain an advantage (instead of for profit).

– *Introduction of a more severe form of extortion committed through actions typical of cybercrimes* (Art. 629, paragraph 2 c.p.).

– *Abolition of the offense of unauthorized possession, dissemination, and installation of equipment, devices, or software intended to damage or interrupt an IT or telematics system* (Art. 615-quinquies c.p.) and simultaneous introduction of a new crime of unauthorized possession, dissemination, and installation of equipment, devices, or software aimed at damaging or interrupting an IT or telematics system (Art. 635-quater, paragraph 1 c.p.).

– *Extension of the crime of damaging IT or telematics systems to those of public interest (instead of public utility)* (Article 635-quinquies c.p.).

– *Increase in penalties for offenses under* Art. 615-ter, 615-quater, 617-bis, 617-quinquies, 635-bis, 635-ter, and 635-quater c.p.

– *Introduction of mitigating circumstances for minor offenses and for those who act to prevent further criminal consequences (*Art. 623-quater and 639-ter c.p.).

– *Changes to the Code of Criminal Procedure regarding:*

– *Extension of the role of the Public Prosecutor* to the Public Prosecutor at the District Court, in relation to jurisdiction over cybercrimes (Article 51, paragraph 3 quinquies of c.p.);

– *Increase in the maximum duration of preliminary investigations for cybercrimes,* extending it from eighteen months to two years (Article 407, paragraph 2 c.p.p.).

– *Amendments to the D. Lgs. 231/2001.*

4. The new features of the law

The D. Lgs. 231/2001 dates back to 2001, also it has been amended several times.

The rule introduced for the first time in Italian law the liability of entities, companies and associations, even without legal personality. Prior to 2001, criminal liability was only of individuals.

The responsibility defined administrative, is manifested as a real criminal liability, because it provides all the procedural guarantees for the natural person accused.

At a first reading it is clear that the D.Lgs. 231 called a very general framework in its formulation, in which they find detail exclusively the two institu-

tes wich exclude liability: the Model of Organization, Management and Control and the Supervisory Body.

For the rest, these are very general rules which have been filled with content over time by case-law.

The D. Lgs. 231/2001 provides that if a crime is committed, by a principal and/or a subordinate, (c.d. precondition offence), the institution or company will also be liable for this, where the conduct of the individual has been committed in the interest and/or for the benefit of that institution or company.

The company will not be liable if – executed risk assessment activity – it has adopted the Organization Model, Management and Control and the Supervisory Body.

In 2001, only computer fraud was envisaged ad a prerequisite for the company's responsibility[5].

The situation has now changed profoundly, since the new DDL Cybersecurity has interacted (also) with the D. Lgs. 231/2001, introducing new types of cases that have become crime-presupposed the responsibility of the company, and that is:

– abusive access to computer systems, Art. 615 ter c.p.

– possession and unlawful distribution of access codes, art. 615 quater c.p.

– illegal interception of computer communications, Art. 617 c c.p.

– damage to public utility computer systems, art 635 quinquies c.p. In addition to the new crimes-assumption[6], what is of interest are the innovations introduced by art. 20 Law 90/2024 to D. Lgs. 231/2001:[7]

– increase of the penalty for in paragraph 1, today between two hundred and seven hundred quotas (first between 100 and 200 shares);

– introduction of the paragraph 1-bis in art. 24 D. Lgs. 231/2001. raising the financial penalty from three hundred to eight hundred euros in relation to the commission of the new case of computer extortion ex art. 629, paragraph 3 c.p.;

– modification of paragraph 2 of art. 24a D. Lgs. 231/2001, increasing the pecuniary sanction up to four hundred euro and that it replaces the crime presupposed, bypassing the art. 615 quinquies c.p. (repealed by art. 16 Law 90/2024) with the new crime referred to in art. 635 quater

[5] https://www.normattiva.it/uri-res/N2Ls?urn:nir:stato:decreto.legislativo:2001-06-08;231.

[6] https://www.gazzettaufficiale.it/eli/id/2024/07/02/24G00108/SG — Legge 28 giugno 2024, n. 90 — Disposizioni in materia di rafforzamento della Cybersicurezza nazionale e di reati informatici, in G.U. Serie Generale n. 153, del 02/07/2024, Art. 16 — Modifiche al Codice penale.

[7] https://www.normattiva.it/uri-res/N2Ls?urn:nir:stato:legge:2024-06-28;90, Art. 20-Modifiche al decreto legislativo 8 giugno 2001, n. 231: 1. All'articolo 24-bis del decreto legislativo 8 giugno 2001, n. 231.

1 c.p., *Possessing, distributing and installing unlawfully equipment, devices or software designed to damage or disrupt a computer or telematic system;*

– integration of paragraph 4 of art. 24-bis D. Lgs. 231/2001: in case of conviction for the crime referred to in paragraph 1 bis, sanctions shall be applied by art. 9 paragraph 2[8], for a period of not less than two years[9].

5. Extortion through computer crime

The only change is the introduction of the crime of extortion through computer crime, also crime-presuppose the responsibility of the institution and company, now provided for in paragraph 3 of art. 629 c.p.: *Any person, by the conduct referred to in Articles 615-ter, 617-quter, 617-sexies, 635-bis, 635-quater and 635-quinquies, or by threatening to do so, compels any one to do or omit anything, If he or someone else makes an unfair profit from the harm done to others, he is liable to imprisonment for a term of six to twelve years and a fine ranging from EUR 5000 to EUR 10000.*

With art. 24-bis D. Lgs. 231/2001, (*Computer crimes and illegal processing of data*)[10],the legislator has established the link between d. lgs. 231 and Cybersecurity.

The element of union is the sharing of the same commitment to protect security, with the adoption of technical and organizational measures capable of reducing cybercrimes committed to the detriment of society.

Art. 24-bis punishes two actions wich are united by the purpose of prevent controls[11].

The inclusion of computer crimes in Decree 231, makes necessary an analysis of the computer risk, to identify protective measures[12].

[8] https://www.normattiva.it/uri-res/N2Ls?urn:nir:stato:decreto.legislativo:2001-06-08;231~art9 — Art. 9 Sanzioni amministrative.

[9] https://www.ambient7.com/wp-content/uploads/2024/07/DDL-cybersecurity-2024_focus-Studio-Legale-Mosetti-Compagnone.pdf, Legge Cyber-security, n. 90/2024, e 231: la responsabilità dell'Ente alla luce della nuova disciplina dei reati informatici.

[10] https://www.normattiva.it/uri-res/N2Ls?urn:nir:stato:decreto.legislativo:2001-06-08;231~art24bis-com3.

[11] https://temi.camera.it/leg18/temi/d-l.105.2019-perimetro.di.sicurezza.nazionale.cibernetica.html #:~:text=Il%20decreto%2Dlegge%2021%20settembre,nei%20settori%20di%20rilevanza%20strategica.

[12] Art. 7 Comma 4.

6. The Method

The method used for corporate liability and Cybersecurity must be the same, based on risk assessment, with particular attention to c.d. relevant processes, with control devices by the Supervisory Body.

The law on Cybersecurity has introduced the mandatory requirement of a preliminary assessment (also) of the cyber risk, according to the proper method already described by D. Lgs. 231/2001.

If this is true as it is true, the assumption rekindles the long-standing question – never solved – about the obligation for institutions, companies and associations to equip themselves with the c.d. Model 231.

Looking through the legislation and integrating its reading with the Explanatory Report[13], the legislator never imposes as mandatory the adoption of the Model of Organization, Management and Control and the Supervisory Body.

But if the new DDL Cybersecurity describes as mandatory the assessment (also) of the cyber risk and if the Law 90/2024 requires that this assessment be performed by the method described in D. Lgs. 231/2001 (not mandatory according to the 2001 Law), how do you resolve the contrast?

Already, before Cybersecurity, the D. Lgs. 231 was considered mandatory, thanks to a complementary reading with other rules dealing with the same subject.

The purpose of D. Lgs. 231 is to allow the adoption of a model aimed at preventing the risk of crime, a goal already adopted by the ordinary legislator.

Art. 2086, paragraph 2, c.c. (Management of the company), as amended by D. Lgs. 14/2019 (c.d. Company Crisis Code) establishes the obligation of the businessman to adopt organizational, administrative and accounting structures adapted to the nature and size of the company also, but not only, on the basis of the timely detection of the company crisis.

He and to act without delay for the adoption and implementation of one of the instruments provided by the order to overcome the crisis and recover business continuity.

In addition, art. 2381 comma 5 c.c., which requires the adoption of the organizational structure (*The delegated bodies shall ensure that the organisational, administrative and accounting structure is appropriate to the nature and size of the undertaking and report to the board of directors and the supervisory board, at intervals fixed by the statutes and in any case at least every six months, the general trend of management and its foreseeable development, as well as operations of major importance in terms of their size or characteristics carried out by the company and its subsidiaries).*

[13] https://www.governo.it/sites/governo.it/files/relazione_illustrativa_86.pdf.

Also, the art. 2475 c.c. has also extended to Limited Liability Companies the provisions whereby the Chairman of the Board of Directors and the CEOs have the obligation to report to the Board of Directors (and to the Statutory Auditors where present), at least every six months, "on the general trend of management" concerning the obligation to adopt appropriate arrangements.

Now, in order to introduce an adequate and effective organizational structure, the entrepreneur will be obliged to carry out a preliminary risk assessment activity, aimed at assessing ex ante the hypotheses of crime that are most likely to occur, Adopting the technical and organizational corrections necessary, according to the only codified method, which is that described in D. Lgs. 231/2001.

If in the opinion of the author, the compulsory liability of companies seemed mandatory even in the past, it is certainly mandatory today, with the Cybersecurity law.

All the subjects addressed by Law 90/2024 will have the obligation to equip themselves, and/ or update where already existing, the Model of Organization, Management and Control and the Supervisory Body.

7. Conclusions

In conclusion, while appreciating the legislative effort of the legislator to have wanted to regulate a sector and a subject so constantly changing, it is perhaps not able to seize the opportunity to harmonize and coordinate a set of rules (Cybersecurity, corporate responsibility, privacy, anti-corruption and transparency regulations, to name just a few) that, from several points of view, all have the same purpose, that of corporate compliance.

SITOGRAPHY

La Relazione annuale al Parlamento, sulle attività svolte dall'Agenzia per la Cybersicurezza nazionale (ACN) in materia di Cybersicurezza, 2024: Acn.gov.it/portale/relazione-annuale

Comunicato stampa del Consiglio dei ministri n. 66 del 25 gennaio 2024: https://www.governo.it/articolo/comunicato-stampa-del-consiglio-dei-ministri-n-66/24831

https://www.4clegal.com

EDMUNDO GRASSI *

The Algomorphic Society: Towards a Theory of Algorithmically Mediated Identities devices§

Abstract

This chapter introduces the notion of algomorphic society to describe a social environment reshaped by algoagents-algorithmic entities that learn, adapt, and act within human worlds. Moving beyond the idea of a merely "digital self," it argues that subjectivity is undergoing a morphological mutation, as algorithmic systems intervene in biographical trajectories, affective economies, and regimes of visibility. Drawing on Simondon's theory of individuation and Malabou's concept of plasticity, identity is framed as a morphogenetic, distributed process co-produced by technosocial affordances and iterative human-machine interactions. Algoagents are analysed as narrative, epistemic, and affective co-actors that structure horizons of expectation and make contemporary selves structurally algo-dependent. The chapter concludes by outlining an algomorphic sociology that displaces methodological anthropocentrism and treats algorithmic infrastructures not as neutral tools, but as actors participating in the production of social reality.

Il capitolo introduce la nozione di società algomorfica per descrivere un ambiente sociale rimodellato dagli algoagenti, entità algoritmiche che apprendono, si adattano e agiscono all'interno dei mondi umani. Superando l'idea di un semplice "sé digitale", si sostiene che la soggettività stia attraversando una mutazione morfologica, poiché i sistemi algoritmici intervengono nelle traiettorie biografiche, nelle economie affettive e nei regimi di visibilità.

* Assistant professor (RTT) in Sociology at the Università Telematica San Raffaele, Roma.

§ This contribution develops the presentation at the conference 'Cybersecurity e prodotti: sfide e responsabilità nelle società interconnesse', 29 November 2024, Italian Supreme Court.

Richiamandosi alla teoria dell'individuazione di Simondon e al concetto di plasticità di Malabou, l'identità viene interpretata come un processo morfogenetico e distribuito, co-prodotto dalle affordance tecnosociali e dalle interazioni iterative uomo-macchina. Gli algoagenti sono analizzati come co-attori narrativi, epistemici e affettivi che strutturano gli orizzonti di aspettativa e rendono i sé contemporanei strutturalmente algoritmo-dipendenti. Il capitolo si conclude delineando una sociologia algomorfica che supera l'antropocentrismo metodologico e considera le infrastrutture algoritmiche non come strumenti neutri, ma come attori partecipi della produzione della realtà sociale.

1. From the Digital Self to the Algomorphic Self

The history of human evolution has always been marked by the trajectory of technoscientific development, which profoundly transforms the processes through which individual identity is constructed. In the digital realm — ranging from the first avatars in virtual worlds to the multifaceted management of social media profiles — the digital self has emerged as an object of narration, negotiation, and visibility. However, the current pervasiveness of algorithms marks a significant rupture in this trajectory: it is no longer a mere extension of the self online, but rather a morphological mutation that affects the very foundations of subjectivity. Algorithms do not simply facilitate or mediate interaction; they analyze data, predict behaviors, guide choices, and intervene directly in shaping both individual and collective biographical trajectories. Within this emerging scenario, I propose the concept of the algomorphic society (Grassi, 2024) to describe a social environment deeply reconfigured by the immanent and semi-autonomous presence of algoagents: algorithmic entities capable of learning, acting, and coexisting with humans within complex relational ecosystems.

Unlike intelligences designed according to an imitative model of the human, algoagents are defined by an operational capacity that does not require the presence of a person. They are grounded in the ability to adapt to novel contexts and to continuously rework experience. The aspiration to model Artificial Intelligence in the image and likeness of the human being reveals itself as a theoretical and cognitive limitation: what is needed instead is the recognition of the emergence of alien intelligences — similar but not identical — which already coexist within the social body, operating as thresholds for the reconfiguration of reality.

The progressive penetration of technoscience into the contemporary cultural fabric has produced a profound configurational reprogramming of individuals, according to alternative relational and grammatical logics. Far from being mere prostheses for enhancement, technological devices now function as co-evolutionary matrices of subjectivity, shaping interdependent identities, augmented personalities, and subjectivities distributed across physical and virtual environments.

Automated and algorithmic forms of self-construction are emerging, in which digital entities no longer merely serve as passive infrastructures (networks, codes, platforms), but act as agents capable of penetrating the most intimate dimensions of experience: affectivity, desire, memory, and decision-making.

This liminal threshold — where the human and the artificial are mutually implicated — gradually dissolves dichotomous boundaries between subject and object, nature and technology, interiority and exteriority. It is here that the ontological co-implication between subjectivity — both human and non-human — and the techno-ecosystem becomes fully manifest, configuring an identity regime based on iterative interactions, continuous modulations, and algorithmic rewritings.

And yet, this new condition is marked by a profound epistemological opacity: we do not know precisely what information is contained within computational models, nor can we fully grasp what algorithms know about us, the world, and our behaviors. Reliable analytical tools to monitor and govern the autonomous evolution of intelligent systems are still lacking. This gap is not merely technical — it strikes at the very heart of the relationship between knowledge, power, and subjectivity in the age of Artificial Intelligence.

2. An Identity Morphogenesis

In the context of the algomorphic society, subjectivity can no longer be conceived as an internal, stable, and coherent structure; rather, it takes on the traits of a relational, mutable, and situated function. It is constituted through distributed interactions between human and non-human agents, within an ongoing process of morphogenesis involving technical devices, informational environments, and cultural codes. In this scenario, identity is neither given nor freely chosen: it is shaped, prefigured, and renegotiated by a set of techno-social affordances that guide the perception of the self and of others. Simondon's (1958) theories on technical individuation offer a fertile framework for

understanding this process: technology is not a mere instrument — it constitutes a mode of existence, capable of transforming both the individual and the collective. In a similar vein, Malabou (2008) identifies plasticity as a principle of subjective configuration and transformation, highlighting the brain's (and the subject's) capacity to rewrite itself in response to received stimuli. In the algomorphic society, such plasticity is amplified by algorithmic presence: each interaction leaves a trace that contributes to defining the parameters of the next, generating an iterative loop of co-modelling between the self and the computational environment.

In this process, algoagents do not merely act as invisible mediators — they fully assume an agentic role (Grassi, 2025). Their agency unfolds along four fundamental dimensions:

— narrative: they construct implicit narratives of the self, proposing personalized trajectories (of consumption, relationships, information);

— epistemological: they filter, select, and rank content, thereby influencing the construction of knowledge;

— affective: they modulate emotional experience by suggesting stimuli or confirmations through sensitive interfaces and reward mechanisms;

— relational: they intervene in the formation of social bonds, acting as catalysts or mediators of interaction.

Algorithmic agency, in this sense, is not limited to the mere execution of commands: it structures the horizon of the possible. Algoagents anticipate needs, define priorities, and determine visibility. Contemporary subjectivity thus becomes algo-dependent — not in a pathological, but in a structural sense: the very form of the self is tied to the continuous operation of systems that process data, recognize patterns, and assign value.

This transformation calls sociology to rethink the categories through which it has so far described identity, interaction, and mediation. We are no longer merely subjects who use technology, but hybrid configurations generated through continuous interaction with computational agents. Subjectivity, in this view, becomes a field of forces: a temporary outcome of negotiation between human desires, algorithmic parameters, and digital environments.

3. Reframing the Social Beyond Anthropocentrism

The algomorphic paradigm demands a structural revision of the framework through which we conceive subjectivity and the social. While sociology has traditionally grounded its analyses on a clear distinction between

human subjects and technical objects, this dichotomy is now proving theoretically inadequate, challenged by emerging currents of thought. The increasing operational autonomy of algoagents, combined with their capacity to interact, learn, and intervene in social dynamics, calls for a rethinking of the very ontology of agency.

Within the algomorphic society, social relations are no longer confined to the human domain; they include — and indeed depend on — non-human entities that function as co-actors. Digital platforms, algorithmic architectures, and intelligent systems are not merely contexts for social action but active components in its genesis and evolution. This realization leads to a reformulation of the concept of interaction, which can no longer presuppose an anthropocentric symmetry: the "other" may be non-human, without diminishing its social relevance.

Algomorphic sociology thus presents itself as a theoretical approach capable of accommodating this transformation. Going beyond the human-technological divide, it treats the algorithmic not as a mere variable to be analyzed, but as a generative element, capable of reorganizing processes of interaction, power dynamics, forms of recognition, and structures of meaning. Algoagents, as semi-autonomous actors, contribute to the co-production of social reality, reshaping cognitive, affective, and normative spaces.

In this scenario, the social appears as a hybrid environment, traversed by data flows, symbolic mediations, predictive calculations, and nonlinear interactions. Identity, embodiment, emotion, and language no longer reside exclusively within the human, but are distributed across a network of relationships that includes algorithmic infrastructures, computational memories, and Artificial Intelligences. What emerges is a transspecies condition, in which coexistence among heterogeneous agents demands a renewed perspective on subjectivity, responsibility, and the modalities of living together. It is no longer a matter of analyzing the impact of technology on society, but of recognizing that society itself is being configured through diffuse algorithmic logics, capable of generating unprecedented forms of subjectivity, agency, and coexistence.

BIBLIOGRAPHY

Grassi E., *Per una sociologia algomorfica: il ruolo degli algoritmi nei mutamenti sociali*, FrancoAngeli, Milano, 2024.

Grassi E., The silent revolution of algoagents: relational reconfigurations and emerging paradigms, in «Sicurezza e Scienze Sociali», II (2025), pp. 69-79.

Malabou C., *What Should We Do with Our Brain?*, Fordham University Press, New York City, 2008.

Simondon G., *Du mode d'existence des objets techniques*, Aubier, Paris, 1958.

MICHELE IASELLI*

Child Safety Online: How AI Can Be Useful and What Are the Critical Issues?§

Abstract

Abstract: The paper examines the role of Artificial Intelligence in protecting minors online, highlighting its potential in preventing cyberbullying, grooming, and exposure to harmful content. It analyzes key AI applications, including content monitoring, safety personalization, and online predator detection. The study also addresses critical issues related to privacy, data protection, algorithmic transparency, and the reliability of automated decisions.

Special attention is paid to risks of excessive surveillance and discrimination. The paper evaluates the contribution of the European AI Act as a regulatory framework for high-risk AI systems. Finally, it explores the emerging legal responsibilities of AI developers and digital platforms.

Il contributo analizza il ruolo dell'Intelligenza Artificiale nella tutela dei minori online, evidenziandone il potenziale nella prevenzione di cyberbullismo, grooming ed esposizione a contenuti dannosi. Vengono esaminate le principali applicazioni dell'AI, come il monitoraggio dei contenuti, la personalizzazione della sicurezza e il rilevamento dei predatori online.

Si approfondiscono, inoltre, le criticità legate a privacy, protezione dei dati, trasparenza algoritmica e affidabilità delle decisioni automatizzate. Particolare attenzione è dedicata ai rischi di sorveglianza e discriminazione. Il contributo esamina l'impatto dell'AI Act europeo come quadro regolatorio per i sistemi ad alto rischio. Infine, si analizzano le nuove responsabilità giuridiche di sviluppatori e piattaforme digitali.

* Docente di diritto digitale e tutela dei dati alla LUISS e di informatica giuridica all'Università di Cassin.

§ This contribution develops the presentation at the conference 'Cybersecurity e prodotti: sfide e responsabilità nelle società interconnesse',29 November 2024, Italian Supreme Court.

1. AI's Contribution

In this context, Artificial Intelligence has established itself as a fundamental resource for preventing and mitigating the risks associated with minors' online presence. Through monitoring online interactions and behaviors, AI can identify dangerous or inappropriate content, as well as detect potential threats in real time.

• Content Monitoring and Analysis

One of the main uses of AI is in monitoring and analyzing content to which minors are exposed online. Social and entertainment platforms, such as YouTube, TikTok, and Instagram, generate an enormous amount of multimedia data that includes images, videos, and texts. AI uses advanced natural language processing (NLP) and image recognition techniques to identify content that could be inappropriate or harmful to minors.

For example, AI algorithms are able to analyze in real time comments left under social media posts, recognizing terms or phrases associated with bullying, hate speech, or grooming. These algorithms are capable of identifying not only individual words, but also the context in which they are used, distinguishing, for example, between a sarcastic comment and a real threat. When suspicious content is detected, AI can immediately alert a human moderator, or in more serious cases, automatically block access to such content to prevent immediate harm[1].

A concrete example of this application is found in the system developed by Google for its YouTube platform, where AI analyzes videos to identify dangerous content, such as violence or inappropriate language. In many cases, AI is able to remove content unsuitable for minors before it can be viewed by a wide audience. This system has made it possible to protect millions of minors from harmful content, ranging from explicit violence to online scams or hate incitement.

• Cyberbullying Prevention

Cyberbullying represents one of the most common and devastating threats that minors face online. This phenomenon occurs when one or more individuals use digital means to intimidate, harass, or humiliate other people, often with serious psychological consequences for victims. AI, through online behavior analysis, can identify bullying patterns in early stages and intervene quickly to prevent their escalation.

An effective example of this application is Instagram's moderation system, which uses Artificial Intelligence to automatically detect bullying messages in private conversations and comments. If AI detects a potentially of-

[1] R. Calo, A. Froomkin, L. Kerr (eds.), *Robot law*, Edward Elgar, 2016.

fensive or threatening message, it intervenes by suggesting the author reflect before sending the message, or by warning recipients of the possibility of reporting the conversation. This technology allows limiting the spread of harmful content before it has negative effects on minors.

Furthermore, many schools and educational organizations are experimenting with AI-based monitoring software to detect signs of cyberbullying among students. These systems are able to analyze students' digital communications and report suspicious behaviors to teachers or school counselors. For example, Bark software, used in several schools in the United States, monitors students' online activities, such as email messages, and immediately alerts educators of any episodes of bullying or other threats, enabling timely intervention.

- Online Predator Detection

Another important field of AI application in protecting minors concerns the detection and prevention of online grooming by digital predators. These individuals often use chats, social media, and online games to gain minors' trust with the intent to exploit them. Thanks to conversation and digital behavior analysis, AI can identify grooming signals, such as suspicious requests for personal information or attempts at psychological manipulation.

A relevant example is represented by AI chatbots designed to simulate minors' behavior in online chats, with the purpose of attracting potential predators. These systems, used by police forces and child protection organizations, allow identifying and neutralizing online predators before they can approach minors. A well-known project is Sweetie 2.0, an Artificial Intelligence developed by Terre des Hommes organization, which simulates the presence of a girl in chat rooms frequented by adults with malicious intentions. The system has already led to the identification and arrest of numerous suspects worldwide.

- Safety Personalization

In addition to preventing immediate threats, AI can also contribute to minors' online safety through safety settings personalization. AI systems are able to adapt to the age, maturity, and behavior of individual minors, dynamically modifying protection levels based on each user's specific needs. For example, inappropriate content filters can be automatically adjusted to become more strict for younger minors and less restrictive as they grow and acquire greater digital awareness.

This personalization also extends to parental control systems, which use AI to monitor minors' browsing habits and provide parents with detailed information on how their children use the Internet. Solutions like Qustodio and Net Nanny use AI to offer in-depth analysis of online activities, suggesting parents modify safety settings based on behavioral patterns. These systems allow blocking po-

tentially harmful content, monitoring device usage times, and receiving real-time alerts when suspicious or dangerous behaviors are detected.

2. Critical Issues Related to AI Use

However, the use of Artificial Intelligence (AI) for protecting minors online, while promising, also presents a series of complex problems that require in-depth analysis. These problems range from issues related to personal data protection and privacy, to those concerning algorithm transparency, automated decision reliability, and operators' legal responsibility. Below, the main criticalities that AI use presents in this context will be examined in depth.

Personal Data Protection and Privacy

One of the main concerns related to the use of Artificial Intelligence in protecting minors online regards personal data protection, particularly in the context of an increasingly stringent regulatory framework like that introduced by GDPR. AI, to function effectively, needs to collect, analyze, and process large amounts of personal data, including those related to minors' online behaviors and interactions. However, processing personal data of vulnerable subjects, such as minors, carries significant risks, both in terms of privacy violation and possible abuse.

GDPR provides very strict rules for protecting minors' personal data, requiring that processing of such data occurs with explicit consent from parents or legal guardians. Furthermore, data collection must respect the principles of minimization and proportionality: only data that is strictly necessary to achieve the purpose should be collected and stored. However, AI, in its continuous monitoring function, could potentially violate these principles by collecting more data than strictly necessary. For example, analyzing a minor's online behavior to detect possible cyberbullying episodes could lead to collecting data on conversations and interactions not relevant from a safety perspective, exposing minors to an unjustified surveillance risk.

Moreover, the risk that these technologies could be exploited inappropriately, for example for commercial purposes, is real. Companies that develop and manage these systems could use collected data to profile minors for advertising purposes, violating their rights to privacy and data protection. This raises important questions about the need to ensure that data collected by AI is used exclusively for protection purposes and not for commercial or other purposes.

Algorithm Transparency and Comprehensibility

Another fundamental problem concerns the transparency of algorithms used by Artificial Intelligence. AI algorithms often operate as "black boxes," i.e., they perform analyses and decisions without their internal processes being comprehensible to non-experts. This raises important issues both ethically and legally, especially in the context of minor protection, where it is essential that parents, educators, and authorities understand how and why certain decisions are made.

The principle of transparency is one of GDPR's pillars and requires that individuals be informed about how their data is processed and for what purposes. However, when an AI algorithm makes a decision – for example, blocking content considered inappropriate or reporting a suspicious conversation – it is not always clear what the decision-making process was. This creates problems not only for user comprehensibility, but also for developers' and operators' responsibility for such systems[2].

Algorithm explainability is an emerging field that aims to make AI decision-making mechanisms comprehensible. However, explainability is often difficult to achieve, especially for more complex machine learning models, such as deep neural networks. Without a clear understanding of how these algorithms work, it becomes complicated for operators to defend or justify decisions made by AI in case of errors or legal disputes. For example, if an AI system erroneously labels an innocent communication between minors as potentially dangerous, it could be difficult for parents or authorities to intervene promptly and correct the error[3].

Reliability and False Positives/Negatives

The reliability of AI algorithms is another crucial element. Although AI technologies are able to analyze and detect large amounts of data with precision and speed that far exceed human capabilities, they are not infallible. False positives (when innocent behavior is labeled as dangerous) and false negatives (when dangerous behavior is not detected) are common problems in AI applications, especially in complex contexts like online minor protection.

A false positive, for example, could occur when an algorithm erroneously interprets a discussion between peers as a cyberbullying episode, while it is actually an innocent conversation. This type of error can lead to unnecessary interventions by parents or authorities, causing unjustified anxiety for minors and their families. On the other hand, a false negative could result in

[2] N. Helberger, J. Pierson, T. Poell, *Governing online platforms: From contested to cooperative responsibility,* in «The Information Society», I (2018), pp. 1-14.

[3] R. Binns, *On the apparent conflict between fairness and explainability in AI,* in «Ethics and Information Technology»,XXII ,I (2020), pp. 1-12.

failing to detect a real threat, such as an online predator attempting to groom a minor, with potentially serious consequences.

To minimize these errors, AI systems must be constantly trained and improved, using always updated and representative datasets. However, the very nature of human behaviors and online content makes it difficult to guarantee 100% accuracy. Furthermore, there is a risk that biases present in data used to train algorithms are reflected in final results, with the risk of discrimination against certain user categories[4].

Legal Responsibility and Regulatory Compliance

The use of Artificial Intelligence also raises important legal responsibility issues that will be examined in depth later. In the context of online minor protection, who is responsible when an AI system fails to detect a threat or, conversely, intervenes inappropriately? This question does not yet have a clear answer in the current legal landscape, but it is evident that responsibility falls primarily on AI system operators and developers[5].

GDPR and other European regulations, such as the Artificial Intelligence Act, seek to establish a clear framework for responsibility in AI system use, imposing transparency, reliability, and safety obligations. However, there is still much uncertainty about how these regulations will be applied in practice, especially when dealing with complex and rapidly evolving technologies like AI. AI system developers must be able to demonstrate that their algorithms have been designed and trained in compliance with legal standards and that adequate control mechanisms have been implemented to prevent harm or abuse[6].

Psychosocial Impact on Minor-Technology Interaction

An often overlooked, but no less important aspect, is the psychosocial impact that the use of Artificial Intelligence-based technologies can have on minors. The idea that their online activities are constantly monitored by automated systems could influence their perception of privacy and their digital behavior. While supervision and protection are essential for safety, there is a risk that overly invasive monitoring could limit freedom of expression and autonomous exploration of the digital world by minors.

[4] S.M. West, *System Error: Where Big Tech Went Wrong and How We Can Reboot*, Random House, 2021.

[5] C. Cath, *Artificial Intelligence and the Good Society: The US, EU, and UK Approach*, Palgrave Macmillan, 2020.

[6] U. Gasser, V. A.F. Almeida, *Human-Centered Artificial Intelligence: Responsible Innovation in the Age of AI,* MIT Press, 2021.

Furthermore, there is concern that excessive reliance on technology could reduce the active involvement of parents and educators in supervising minors' online activities. If parents rely completely on AI to protect their children, they might neglect the need to educate them about digital world risks and teach them to navigate safely and securely.

3. The AI Act Contribution

The AI Act (Regulation No. 1689/2024) recently enacted by the European Union represents one of the most important regulatory interventions regarding Artificial Intelligence and has a significant impact on the context of online minor protection. The law was conceived to establish a clear and harmonized legal framework for AI use in all sectors, with particular attention to systems that can influence individuals' fundamental rights, including minors' safety, health, and privacy.

One of the AI Act's key elements is the classification of AI systems based on the risk level they present for fundamental rights and freedoms. In this context, AI systems used for online minor protection, such as those employed to monitor content or detect suspicious behaviors, are considered high-risk systems. This classification requires companies and organizations using such systems to adopt rigorous compliance measures to ensure that AI operates safely, transparently, and responsibly.

The AI Act considers that all high-risk AI systems must be subject to impact assessments, with particular reference to risks related to personal data protection and discrimination. In the case of minor protection, this means that any system used to monitor content or prevent online grooming must be carefully designed to avoid unwanted side effects, such as excessive surveillance or invasive profiling. In particular, the AI Act provides that AI systems using biometric data, such as facial or voice recognition to identify potential threats, must undergo thorough evaluation to ensure privacy rights compliance[7].

One of the AI Act's main objectives is to improve the transparency of algorithms used in Artificial Intelligence systems, especially when these have a direct impact on people's rights. In the context of online minor protection, transparency is a crucial aspect. Parents and legal guardians must be able to understand how systems that monitor their children's online activity work and on what basis decisions are made such as blocking certain content or sending alerts to competent authorities.

[7] B. Goodman, S. Flaxman, *European Union Regulations on Algorithmic Decision-Making and a "Right to Explanation"*, in «AI», 38, III (2017), pp. 50-57.

The AI Act introduces specific transparency obligations, requiring AI developers to provide detailed information about algorithm functioning, including ways personal data is processed and logics behind automated decisions. This is particularly relevant for online minor protection, as transparency helps create trust between parents and institutions and allows more rigorous control over decision-making processes. For example, an AI system that reports a potential threat or dangerous behavior must be able to explain clearly and comprehensibly which factors led to that conclusion, avoiding such decisions being perceived as arbitrary or discriminatory.

Furthermore, the AI Act provides that AI system users – in this case, parents or guardians – have the right to receive information about safety measures implemented to prevent errors or abuse. This is a fundamental step toward greater responsibility of AI developers and operators, as it requires them to make internal mechanisms of systems managing such delicate data accessible and comprehensible.

Another essential contribution of the AI Act is the introduction of specific obligations regarding fundamental rights impact assessment. This is a key element for AI systems involving minors, as protecting their rights, including those related to privacy, education, and safety, is considered priority[8].

In practice, the AI Act requires organizations to perform an impact assessment before implementing high-risk AI systems. This assessment must identify potential risks to fundamental rights and freedoms and propose measures to mitigate such risks. In the case of online minor protection, this means that AI systems must be designed so as not to unduly compromise minors' right to privacy or their access to legitimate educational or recreational content.

The impact assessment required by the AI Act must also consider the possibility of discrimination. A common problem in AI algorithms, in fact, is the tendency to replicate biases present in training data. For example, if an AI system has been trained on a limited or biased dataset, it might erroneously classify some behaviors or content as inappropriate for some user groups, potentially discriminating against minors from certain ethnic or cultural backgrounds. The AI Act seeks to prevent this problem by requiring that data used to train algorithms be representative and free from bias, and that control mechanisms be implemented to identify and correct any errors.

The AI Act requires that high-risk AI systems be subject to constant monitoring and supervision, both by operators and competent authorities. In the context of online minor protection, this means that automatic content mo-

[8] J. Fjeld, N. Achten, H. Hilligoss, A.C. Nagy, M. Srikumar, *Principled artificial intelligence: Mapping consensus in ethical and rights-based approaches to principles for AI*, Berkman Klein Center Research Publication No. 2020-1, 2020.

nitoring systems must be periodically checked to ensure their effectiveness and compliance with regulations.

In particular, the AI Act introduces the obligation for companies to implement human review mechanisms. This aspect is particularly important in case of errors or discrepancies in automated decision-making processes. For example, if an AI system erroneously detects content as dangerous and blocks viewing, or if it fails to detect a real threat, it must be possible for human operators to intervene to review and correct such decisions. This principle, known as meaningful human intervention, is fundamental to prevent AI from acting completely autonomously without any control.

In addition to internal monitoring, the AI Act provides that national regulatory authorities have the task of supervising AI system use. This introduces an additional level of control, ensuring that systems used for minor protection comply with safety and transparency standards defined by law. Regulatory authorities can also require periodic audits to verify that AI systems have been implemented correctly and that there are no hidden risks for minors.

4. Responsibilities Arising from AI Use

The introduction of AI in the context of online minor protection has inevitably created new forms of responsibility, which extend across different levels and involve developers, service providers, and digital platforms. These emerge regarding the adoption of AI systems to monitor, analyze, and prevent potentially harmful behaviors for minors, such as cyberbullying, online grooming, and exposure to inappropriate content. In an increasingly stringent regulatory environment, such as that imposed by the AI Act and GDPR, responsibility obligations expand not only to those who implement such technologies, but also to those who develop and manage them.

One of the central aspects of new responsibility concerns AI system creators. These subjects must ensure that algorithms used for minor protection comply with principles of transparency, fairness, and non-discrimination. Unlike traditional software, AI bases its decisions on machine learning processes, which require analysis of large amounts of personal data, including those related to minors' online behaviors. This requires particular attention to privacy protection and minimization of collected data. Developers, therefore, are legally responsible for integrating "privacy by design" and "privacy by default" principles, established by GDPR, into their system design phases to

ensure that developed technologies protect personal data at every processing stage. In case of violations, such as excessive data collection or processing, developers can be held responsible for harmful consequences, especially if data involves minors, who are particularly vulnerable subjects.

Furthermore, the AI Act introduces a new level of responsibility that requires developers and providers of high-risk AI systems, such as those used for minor protection, to perform impact assessments before marketing or implementing their products. These assessments must highlight potential risks that AI use could pose to fundamental rights, particularly regarding minors' privacy, safety, and non-discrimination. A system that does not adequately consider these factors or that presents security flaws could easily be exposed to criticalities that, if not managed in time, could result in serious legal responsibilities for developers.

Platforms using these AI systems are also involved in this responsibility chain. When a digital platform adopts an AI system to monitor and manage content that may be harmful to minors, it is essential that such technology be managed transparently and in compliance with current regulations. If an AI algorithm erroneously identifies safe content as dangerous, or fails to detect a real threat promptly, the platform can be called to answer for damages suffered by involved minors. These platforms' responsibility is particularly critical, given that they work as digital intermediaries that facilitate minors' interaction with the online world. They must therefore ensure that AI systems used for protection do not compromise users' rights and that automated decisions are subject to human review when necessary. This, as seen, is in line with the principle of meaningful human intervention.

Legal responsibilities also arise from potential damage that could derive from inappropriate or ineffective use of AI systems. If an AI technology does not adequately detect an online grooming episode or if it unjustifiably blocks legitimate content for minors, consequences could be serious both for victims of these errors and for involved platforms. Companies implementing such systems could face claims and sanctions, especially if they have not adopted adequate preventive measures to ensure their system's safety and transparency. Non-compliance with regulations imposed by the AI Act or GDPR can result in high fines and legal proceedings, especially in a context like minor protection, where damage can have long-term impact on involved subjects' emotional and psychological health.

Another area of responsibility concerns protection against discrimination. AI algorithms are susceptible to reproducing biases present in data they were trained with, which could lead to discrimination against certain groups of minors based on ethnicity, gender, or other socio-demographic factors. For example, if an AI system tends to erroneously classify certain behaviors of a

specific ethnic group as threatening or inappropriate, this could lead to unfair or discriminatory treatment. In such cases, both AI system creators and platforms using them can be held responsible for damage caused by such biases, even if involuntary. As already highlighted, regulations like the AI Act require that AI systems be designed to be fair and bias-free, imposing rigorous tests and regular audits to ensure there is no discrimination in automated decision-making processes.

Finally, new forms of responsibility also emerge in relation to users' right to appeal, including the right to contest automated decisions made by AI algorithms. In cases where a parent or guardian believes that an AI system has unjustly blocked content or allowed harmful behavior, they must have access to appeal mechanisms to request a decision review. Companies are therefore responsible for providing accessible and transparent tools to manage these complaints, ensuring that every automated decision is supported by comprehensible explanations and adequate review processes. This obligation ensures that the AI system does not operate as a "black box," but that its decision-making processes are transparent and accountable.

BIBLIOGRAPHY

Binns R., *On the apparent conflict between fairness and explainability in AI*, in «Ethics and Information Technology», XXII, I (2020), pp. 1-12.

Calo R., Froomkin A., Kerr I. (eds.), *Robot law*, Edward Elgar, Cheltenham–Northampton 2016.

Cath C., *Artificial Intelligence and the Good Society: The US, EU, and UK Approach*, Palgrave Macmillan, London 2020.

Fjeld J., Achten N., Hilligoss H., Nagy A. C., Srikumar M., *Principled artificial intelligence: Mapping consensus in ethical and rights-based approaches to principles for AI*, Berkman Klein Center Research Publication No. 2020-1, Cambridge (MA) 2020.

Gasser U., Almeida V. A. F., *Human-Centered Artificial Intelligence: Responsible Innovation in the Age of AI*, MIT Press, Cambridge (MA) 2021.

Goodman B., Flaxman S., *European Union Regulations on Algorithmic Decision-Making and a "Right to Explanation"*, in «AI», XXXVIII (2017), pp. 50-57.

Helberger N., Pierson J., Poell T., *Governing online platforms: From contested to cooperative responsibility*, in «The Information Society», XXXIV (2018), pp. 1-14.

West S. M., *System Error: Where Big Tech Went Wrong and How We Can Reboot*, Random House, New York 2021.

ENRICHETTA GIULIA JANNOTTI PECCI

Blockchain at the service of Cybersecurity in the Italian food and agri-food supply chain[§]

Abstract

This paper examines the role of blockchain technology as an innovative tool supporting Cybersecurity in the Italian agri-food supply chain, with particular attention to food safety, production traceability, and the protection of data and intellectual property. Starting from an analysis of the growing risks of cyberattacks and data breaches affecting the manufacturing and agri-food sectors, the study highlights how the digitalization of value chains exposes businesses to significant vulnerabilities, especially with regard to the theft of strategic information and the phenomena of counterfeiting and the evocation of Italian origin. In this context, blockchain is analysed through its main structural features – decentralization, immutability, transparency, and the use of cryptography – which make it suitable for ensuring data integrity and security throughout the entire supply chain. The paper further explores the practical applications of blockchain in enhancing agri-food traceability, safeguarding the authenticity of Made in Italy products, and protecting innovation and intangible assets within the sector. In conclusion, it is argued that blockchain can represent a valuable tool in support of Cybersecurity policies, as well as food and technological sovereignty, while also highlighting the need for an appropriate balance between technological solutions and legal instruments, together with public policies aimed at fostering its conscious, efficient, and effective adoption.

L' articolo analizza il ruolo della tecnologia blockchain come strumento innovativo a supporto della sicurezza informatica nella filiera agroalimentare italiana, con particolare attenzione alla sicurezza alimentare, alla tracciabilità

[§] This contribution develops the presentation at the conference 'Cybersecurity e prodotti: sfide e responsabilità nelle società interconnesse',29 November 2024, Italian Supreme Court.

della produzione e alla protezione dei dati e della proprietà intellettuale. Partendo da un'analisi dei crescenti rischi di attacchi informatici e violazioni dei dati che colpiscono i settori manifatturiero e agroalimentare, lo studio evidenzia come la digitalizzazione delle catene del valore esponga le imprese a significative vulnerabilità, soprattutto per quanto riguarda il furto di informazioni strategiche e i fenomeni di contraffazione e evocazione dell'origine italiana. In questo contesto, la blockchain viene analizzata attraverso le sue principali caratteristiche strutturali – decentralizzazione, immutabilità, trasparenza e uso della crittografia – che la rendono adatta a garantire l'integrità e la sicurezza dei dati lungo l'intera filiera. Il documento esplora inoltre le applicazioni pratiche della blockchain nel migliorare la tracciabilità agroalimentare, salvaguardare l'autenticità dei prodotti Made in Italy e proteggere l'innovazione e i beni immateriali del settore. In conclusione, si sostiene che la blockchain può rappresentare uno strumento prezioso a sostegno delle politiche di sicurezza informatica, nonché della sovranità alimentare e tecnologica, sottolineando al contempo la necessità di un adeguato equilibrio tra soluzioni tecnologiche e strumenti giuridici, insieme a politiche pubbliche volte a favorirne l'adozione consapevole, efficiente ed efficace.

1. Introduction

This paper aims to identify, from an innovative perspective, ways to protect agri-food supply chains from potential cyber-attacks by leveraging emerging technologies such as blockchain.

The fields of analysis are those related to food safety in the agri-food sector supply chains, as well as the risks deriving from data breaches, in particular with regard to the matter of intellectual property protection.

These fields have a minimum common denominator identifiable in the broad concept of protection of Made in Italy and Italian identity. This observation derives from the circumstance that each sector reported assumes a notable importance in relation to the impact it generates in the Italian social and productive system, and therefore also particularly exposed to the risks of the digitalization of society and markets that accompany the contradictions deriving from the era of hyper-globalization in which the contemporary world lives. If on the one hand there are the material risks of the opening of markets that can damage the supply chains, directly or indirectly, as in the case of counterfeiting activities or evocation of products of the Italian supply chain; on the other hand the digitalization processes have allowed that the threats can be paraly-

zing of the production systems, extortionate, intrusive, counterfeit and also difficult to identify precisely because of the delocalization and dematerialization of technological tools, as well as disturbing the works of genius and intellect.

To understand the threat to businesses and the risk of paralysis resulting from these activities, it is appropriate to report the data collected by the National Cybersecurity Agency during 2023: out of 1,411 cyber events, 38 involved the manufacturing sector, with a global growth of +25% in incidents damaging the manufacturing sector in the four-year period 2019-2023, also recording for the Italian manufacturing sector a high risk level for almost 40% of cases in 2022 and almost 30% in 2023, noting as a further alarming fact that a quarter of the total attacks aimed at manufacturing at a global level concern Italian manufacturing companies[1]. This data should not be overlooked, but rather stimulates further reflection, as it highlights that in the relationship with sectors traditionally victims of cyber attacks, such as telecommunications and public administrations, the Italian manufacturing sector has become one of the main target sectors[2], also defining a worrying picture of the protection capacity of companies, highlighting among other things the evident weakness of defense techniques that are not sufficiently up to the level of criminal attack techniques and that the presence of vulnerabilities makes these targets particularly attractive[3]. In particular, the Clusit 2024 Report reports the TALPA operation relating to the prosecution of the Ragnar Locker group, which hit, among others, several Italian companies in the food sector through ransomware attacks[4].

Coming to the topic of food safety, this represents various facets: first of all, the question of supply chain safety was raised from a health point of view, in light not only of the pandemic crisis that has highlighted the risks deriving from the uncertainty of supply chain transparency and the difficult controllability of production methods in other continents[5-6], but also in a general way by the existence of such a long and complex supply chain whose tracking is often complex and articulated and not easily reconstructable. Furthermore, the risks that arise not only pose an issue of health safety but also of commer-

[1] Clusit Report 2024, pp. 15; 23; 38.

[2] ACN Annual Report 2023, pp. 13-16.

[3] Clusit Report 2024, page 39.

[4] Clusit Report 2024, page 102.

[5] A. Iftekhar, X. Cui, *Blockchain-Based Traceability System That Ensures Food Safety Measures to Protect Consumers Safety and COVID-19 Free Supply Chains*, in «Food», X (2021), p. 2.

[6] A. Iftekhar, X. Cui, *Blockchain-Based Traceability System That Ensures Food Safety Measures to Protect Consumers Safety and COVID-19 Free Supply Chains,* cited, p. 7, in particular, the Authors highlight the inadequacy and inability of traditional traceability methods to ensure supply chain transparency and therefore propose innovative tools such as blockchain.

cial correctness, as occurred in the cases of fraud for the sale of horse meat instead of cow meat in 2013 with reference to the well-known Irish case[7].

Although the topic of food safety is undoubtedly relevant, it nevertheless intersects with a further crucial issue in the reflection, which is that relating to the Cybersecurity of supply chains.

To the question why Cybersecurity concerns the agricultural and agri-food supply chain, it is appropriate to answer starting from the threat that digital technologies can represent for productivity and the correct performance of the same. In particular, the main risk is that of data breaches, which can put farmers, agricultural and agri-food businesses at risk, but above all can attribute to cyber criminals something with a higher value both from a value point of view and from an economic and monetary point of view, which are the information and data relating to agricultural production[8]. These attacks therefore translate into identity theft.

Even more risky and damaging are ransomware, extortion attacks that threaten to destroy all data relating to production in exchange for money. Threats of this nature therefore invite us to consider the specifically national aspects deriving from such damaging activities and on which we want to focus: the Italian territory, if we tried to read it with the eyes of a statistician, would take on the characteristics of a boot full of precious data relating to agricultural and agri-food production, traditional and non-reproducible production methods, centuries-old stories that are handed down, innovative information that tells the potential and technological and economic growth, crucial for an economically and industrially strategic sector. Such data demands protection.

The further aspect on which we want to focus is closely linked to agri-food production, but it can also take on a broader scope and not just limited to this field, and concerns the protection of intellectual property. It plays a crucial role in the growth and competitive advantage of a country, by virtue of the creative and technological value that works of the intellect generate. Even the intellectual properties and related assets have significant flaws and are therefore particularly exposed to pirate activities such as fraud or any other malicious activity.

The protection of intellectual property must not only revolve around adequate legal measures, communication and educational campaigns, but necessarily involves the implementation of adequate Cybersecurity measures that guarantee data protection[9].

[7] A. Iftekhar, X. Cui, M. Hassan, W. Afzal, *Application of blockchain and internet of things tamper-proof data availability for food safety*, in «Journal of Food Quality» , X (2020), pag. 2.

[8] Cybersecurity. Agriculture and food security, 2024, USAID, p. 1.

[9] C. Mavani, H.K. Mistry, R. Pate, A. Goswami, *The Role of cybersecurity in protecting intellectual property,* in «International Journal on Recent and Innovation Trend in Computing and Communication», XII (2024), pp. 529-53.

In particular, the growing pervasiveness of digital technologies is revolutionizing our way of doing business, especially modifying industrial processes and supply chains, requiring a strong commitment to supporting Cybersecurity that requires an open and unbiased reflection in the identification of valid and transparent tools capable of protecting data, supply chains and people, such as blockchain. Therefore, blockchain could be the tool capable of tracking the supply chain in all its phases, from the development of the idea to the marketing of the product, ensuring the protection of the integrity of the legal asset of the supply chain while also ensuring the value of the authenticity of the product; this allows two protection mechanisms, protection of security integrity and protection of the identity of the supply chain to be ensured by a single tool that is blockchain.

2. Why blockchain?

The issue that comes to light concerns the traceability of the product and the traceability of the supply chain and production, in order to guarantee transparency on the one hand but also safety and protection on the other.

The United Nations Global Compact Office defines traceability as "the ability to identify and trace the history, distribution, location, and application of products, parts, and materials, to ensure the reliability of sustainability claims, in the areas of human rights, labor (including health and safety), the environment, and corruption"[10]. This definition is therefore in line with the need of consumers and businesses to identify a tool that can satisfy the transparency of the supply chain, its safety and the minimization of the risk of damage or attack on it.

But why blockchain? What are the characteristics of this technology that allow it to satisfy these multiple needs?

Blockchain is a technology based on a Distributed Ledger Technology, that is, on a system of distributed registers consisting of a database of digital assets in a network formed by a multitude of subjects, called nodes of the network, each of which owns a copy of the entire register, which is updated every time there is an intervention or a variation[11]. Blockchain is a particular species of a genus which is that of DLT.

[10] B.S. Norton, *A guide to Traceability: A practical approaches to Advance Sustainability in Global Supply Chains,* United Nations Global Compact Office, 2014, New York.

[11] A. Gaschi, V. Portale, *The definition of blockchain and Distributed Ledger Technology,* in «Blockchain and smart contracts. Operation, legal and international profiles, practical applications», Milan, Giuffrè, 2019.

The structure of the blockchain can be described by identifying two large groups. The first is that of public blockchains, which contains two subcategories, the permissionless and the permissioned ones, and then the second is that of private blockchains, which only have the permissioned variant. Public blockchains allow anyone to access in a pseudonymous form, without any formality. In public blockchains there are no central authorities to control access or the general functioning of the platform[12].

In private blockchains, users with known identities operate, who must meet certain requirements in order to become nodes in the chain. Therefore, access to these blockchains is permitted only with the consent of the centralized authority in charge that is responsible for monitoring the security of the platform and the reliability of individual members[13]. This type of blockchain, although it sacrifices some of the horizontal characteristics typical of the traditional model, offers greater security guarantees and also allows for advantageous technical effects, such as greater chain update speed[14].

It is easy to understand that the system in question presents highly complex characteristics, however, it is possible to identify a series of characteristics that constitute its distinctive features. The first of these is decentralization, concerning the organizational method[15]. If we look at the blockchain as a database, we must imagine a network of users connected to each other, in conditions of equality, in different places, without the intervention of an external authority: each user of the chain has a copy of the blockchain, identical and of equal authority with respect to the others[16]. In particular, each user has both an active and passive role, since he is both creator and certifier as well as supervisor and guarantor of the legitimacy of the operations[17].

Blockchain, therefore, introduces a new practice to ensure the reliability of the system, which is decentralized and distributed. Trust is placed in the system as a whole, because it is entrusted to the cryptographic verifiability of the chains and not to the individual user; it is, therefore, a "trustless trust" system[18]. However, this does not mean that public blockchains are not equipped with a

[12] P. De Filippi, A. Wright, *Blockchain and the Law Rule of Code*, Harvard University Press, Cambridge, 2018, pp. 38 et seq.

[13] A. Berke, *How safe are blockchains? It depends*, in «Harvard Business Review», (2017).

[14] A. Narayanan, J. Bonneau, E. Felten, A. Miller, S. Goldfeder, *Bitcoin and Cryptocurrency Technologies: A Comprehensive Introduction,* Princeton University Press, 2016, pp. 133 et seq.

[15] P. De Filippi, A. Wright, *Decentralized Blockchain Technology and the Rise of Lex Cryptography,* in «Social Science Research Network», 2015.

[16] P. Cuccuru, *Beyond bitcoin: an early overview on smart contracts*, in «International Journal of Law and Information Technology», III (2017), pp. 182-183.

[17] *Ibid.*

[18] P. De Filippi, A. Wright, *Decentralized Blockchain Technology and the Rise of Lex Cryptographia* , cited.

governance system, but it is an unexpressed form as it manifests itself informally through conglomerates of nodes that act in a harmonized way, although a system similar to that of private blockchains would certainly be preferable, since they would benefit in terms of transparency. The element of decentralization consists in the fact that in both public and private blockchains the distribution of registers takes on a global character, going beyond national borders[19].

Another feature is the immutability of transactions or information recorded on the blockchain. It is important to highlight that there is a distinction between the two operations: in the first case, the transaction (understood in legal terms as any exchange of data, even those inherent to transactions in a legal sense) must be seen as a legal fact whose irreversibility is guaranteed by distributed ledgers, so that what has been done cannot be brought back to the status quo ante ; while in the second case, the blockchain has the nature of a distributed ledger, identifying immutability in the guarantee that the recorded data cannot be manipulated[20]. Furthermore, the immutability data is what makes the ledger certain. However, it is important to clarify that the system does not certify the truthfulness of the information, but only the elements relating to the declaration of the information.

Finally, the last relevant feature is the transparency resulting from the use of peer-to-peer structures and encryption[21].

The blockchain has the hash function as its main computing foundation and demonstrates the importance of cryptography in the functioning of the chain, in particular to ensure the requirement of immutability. The hash function is a mechanism by which input information is converted into output data, consisting of a variable-length string called hash value. This cryptography technique is normally used to verify the integrity of data or a message that could be altered during transmission. However, the function is not a finished product but is the basis for the construction of more complex mechanisms that take the general name of hash chain or also known as Merkle Tree. The Merkle tree is a computer structure that allows you to verify the membership of a specific piece of data to a data structure, which has a linear form. The tree is made up of a diagram of the elements that are verified and then paired and replaced with their hashes. The result of the sum of two hashes, i.e. the inputs, generates a node (an operation that is repeated for each piece of data to be

[19] P. De Filippi, A. Wright, *Blockchain and the Law Rule of Code*, cited, pag. 37.

[20] For ideas on the use of blockchain as a register to maintain data integrity, see E. Gaetani, L. Aniello, R. Baldoni, F. Lombardi, A. Margheri, V. Sassone, *Blockchain-based Database to Ensure Data Integrity* in «Cloud Computing Environments», in «Proceedings of the First Italian Conference on Cybersecurity» (ITASEC17), (2017).

[21] P. De Filippi, A. Wright, *Blockchain and the Law Rule of Code*, 2018, pag. 37.

verified) that leads to the structuring of a tree[22]. Therefore, the blockchain is made up of a set of groups called blocks within which there is information relating to the type of operation performed. Each block is connected to the previous one via a pointer, a programming element that indicates the position of certain data within a program, and which forms the links of the chain in the blockchain[23].

Therefore, the immutability of the chain is determined precisely by the consequentiality of the blocks through the use of the hash value, and even a small change that breaks the correspondence of input and output will be capable of destroying the chain.

The element that protects the blockchain structure is the consensus mechanism, which is called proof of work[24]. This tool is strictly linked from an IT point of view to the hash function, in fact, when a block is added to the chain it is necessary to first guarantee the validity of the hash string. This consensus mechanism is useful to avoid and prevent possible tampering of the chain, or for changes to the chain itself which, however, in order to be changed, the hash function should necessarily be directly affected.

Therefore, this form of consensus is called "distributed" because it is the result of the complex of actors involved in the chain and is not expressed by a central authority. Therefore, it is not only connected to the immutability of the blockchain but is also the foundation of its decentralization in combination with network technologies[25].

Finally, as anticipated, the blockchain ensures its reliability and transparency thanks to encryption and digital signature technology. Encryption is an operation by which the content of a message is not made clear so that it is not known to people who are not authorized to read it.

Encryption is based on a system of keys to encrypt and decrypt a message, so this must be known to both the sender and the recipient[26]. To overcome the problem of passing the key for reading, a double key encryption was developed, one of which is public and one is private. Through this system, it is possible to create a digital signature mechanism that ensures the security of the data entered: each user receives two keys, one to create a digital signature for the work (transition or recorded information), and a

[22] See D. Carboni, *The technologies underlying blockchain, in M. Giordano, R. Battaglini, Blockchain and smart contracts. Operation, legal and international profiles, practical applications,* Milan, Giuffrè, 2019.

[23] R. Ballard, *Blockchain: The facts and Fiction*, in «Oxford Academic Journal», II (2020), pag. 22.

[24] See D. Carboni, *The technologies underlying the blockchain,* cited.

[25] P. De Filippi, A. Wright, *Blockchain and the Law Rule of Code,* cited.

[26] See D. Carboni, *The technologies underlying the blockchain,* cited p. 37.

public one known to all users of the network to verify the digital signature, the identity of the counterparty and the interaction on the chain[27].

Therefore, if we want to summarize the elements that give blockchain the privileged position among the tools at the service of Cybersecurity, they are: confidentiality, offered by functions that aim to guarantee the anonymity of the user; data integrity, the use of the hash function allows the recorded data to be made immutable and therefore protected from attacks aimed at modifying the data; availability, consisting in the ability to always guarantee that the correct information remains accessible to other peers and alive even in the event of a compromise. Therefore, it is the very nature of blockchain that makes it an adequate tool for the preservation of information and data, and for the traceability of the supply chain according to transparency and security logics thanks also to the application of cryptography considered a fundamental node of Cybersecurity.

3. The use of blockchain to guarantee traceability and safety of the food and agri-food supply chain

The purposes of using blockchain can have a dual interpretation: on the one hand, it represents the tool thanks to which it is possible to ensure greater transparency of the supply chain, so as to protect the interest of the consumer, and the need of the producer to be protected from evocative or counterfeit practices carried out by third-party producers; on the other hand, blockchain can become the tool thanks to which it is possible to protect the supply chain from attacks by third parties whose aim is to destroy the production operations. This protection is realized in the functionality of encrypting the information contained in the supply chain so as to prevent those practices of data modification. In fact, the impossibility of modifying the content of the supply chain in blockchain allows to satisfy those needs deriving from the needs of food safety and protection of the supply chain, producers and consumers as a whole, but also of a security nature with regards to the protection of the supply chain from intrusions by third parties.

Supply chains have undergone several shocks in recent years that have put their reliability and proper functioning at risk. In particular, not only has the certainty of global supply chains been called into question by interven-

[27] H. Natarayan, S. Krause, H. Gradstein, *Distributed Ledger Technology (DLT) and Blockchain*, in «World Bank Group FinTech», I (2017), pp. 8 et seq.

tions to rethink production chains, but also the pervasiveness of technology and its degenerations has required a rethinking of the levels of security and protection of value chains. In fact, the centralization model, that is, companies that store information regarding their production and their supply chain in centralized databases[28], presents high risks due to its vulnerability to attacks carried out by malicious entities. Therefore, taking note of the problem connected to the need for traceability of the supply chain, some have believed that blockchain could be a useful tool for companies to be able to store and share information on the life and path of their products in the supply chain, according to a rule of maximum transparency and immutability of data[29], thus proving that the use of systems based on distributed registers that guarantee the immutability and inalterability of data can ensure promising applications in the supply chain[30].

This opportunity has been tested and implemented by various entities. At an international level, the IBM Food trust has been implementing, a system of traceability of the food supply chain based on blockchain and which allows immediate access to data and information relating the production, and adopted by large companies in the food sector such as Nestlè and Carrefour[31].

However, of greater interest is instead to analyze what has been put in place by Italian companies to ensure that there is a strengthening in the use of technologies by promoting a necessary increase in the levels of transparency and security[32].

[28] T. K. Agrawal, V. Kumar, R. Pal, L. Wang, Y. Cheng, *Blockchain-based framework for supply chain traceability : a case example of textile and clothing industry*, in «Computers & Industrial Engineering», (2021), p. 154.

[29] S. Saberi, M. Kouhizadeh, J. Sarkis, L. Shen, *Blockchain technology and its relationships to sustainable supply chain management,* in «International Journal of Production Research», (2019), pp. 2117-2135.

[30] This attention to the use of DLTs in the supply chain is confirmed by the European Parliament Resolution of 3 October 2018 on distributed ledger technologies and blockchain: building trust through disintermediation (2017/2772(RsP), https://eur-lex.europa.eu/legal-content/lt/tXt/PDf/?uri=CeLeX:52018IP0373, point 16: "*underlines the importance of DLT in improving supply chains; notes that DLT can facilitate the shipping and tracking of the origin of goods and their ingredients or components, improving transparency, visibility and compliance control, ensuring that sustainability and human rights protocols are respected at a product's place of origin, thus reducing the risk of illegal goods entering the supply chain and ensuring consumer protection; notes that DLT can be used as a tool to improve the effectiveness of counterfeit controls carried out by customs officials".*

[31] B. Khalkhal, *IBM Blockchain Transparent Supply & Food Trust. Creating Sustainable Value Chains through Blockchain*, IBM.

[32] L. Compagnucci, D. Lepore, F. Spigarelli, E. Frontoni, M. Baldi, L. Di Bernardino, *Uncovering the potential of blockchain in the agri-food supply chain: an interdisciplinary case study,* in «Journal of Engineering and Technology Management», (2022), pag. 8.

This tool also represents an opportunity to enhance the Italian agricultural and agri-food supply chain since, precisely in this area, promotion passes through the implementation of the digitalization and technologization of the same. In fact, as Arisi emphasizes, "the digitalization of the supply chain makes its components legible and allows a more pervasive principle of control over the product"[33], a control that however should not be understood in "stringent or limiting" terms, but rather takes on a positive and protective meaning for both the consumer and the producer, since control should be interpreted as traceability.

This is also confirmed by the guidelines of European agricultural policy, which, recognising the strategic value, with particular attention to the indispensability of protecting food supply chains, encourages the use of digital tools also in the tracking of agri-food products, in light of the economic, environmental and social value of which they are holders and bearers[34-35].

It is therefore possible to affirm that the protection of the Italianness of the supply chain is not only a question of protecting the producer or the consumer, but it is also a question of protecting a legal asset that can be protected by virtue of its territorial, historical and productive impact, as well as its strategic impact on the autonomy of food supply chains.

4. Protecting the agricultural and agri-food supply chain and intellectual property from the risks of data breaches

As reported above, during the four-year period 2019-2023 there has been an increase in information attacks in the Italian manufacturing sector, reaching a world record. This fact highlights the particular value of the data belonging to the national production chains and the data concerning the manufacturing sector as a whole, a value not only of an economic nature but also of a strategic and informative nature with regard to the production and industrial processes implemented by Italian companies. The risks deriving from data breaches concern first of all the risk deriving from the dissemination of sensitive data, such as those concerning intellectual property, works of the intellect and any other invention or peculiarity of manufacturing.

In particular, considering the scope of the food and agri-food supply chain, the use of tools to protect the supply chain from any possible dange-

[33] M. Arisi, *Cultivating Trust? A Legal Perspective on Data, Blockchain, and Traceability Along the Agri-Food Supply Chain,* in «Quaderni della Rivista di diritto alimentare», I (2023), p. 56.

[34] European Commission, Communication COM(2019) 640.

[35] M. Arisi, *Cultivating Trust? A Legal Perspective on Data, Blockchain, and Traceability Along the Agri-Food Supply Chain*, cited, p. 56.

rous and prejudicial action takes on the connotation of protection of a strategic asset in line with policies on food sovereignty understood as security of supplies.

In light of this, production factors play a key role in relation to food sovereignty and it is therefore necessary to take into account the inputs deriving from the agricultural sector in terms of innovation. In fact, scientific and technological progress in the varietal and agricultural fields, consisting in the advancement in terms of digitalization of the agri-food chain, is capable of contributing to the promotion of food security and food safety objectives, essential elements for the pursuit of food sovereignty[36].

Therefore, today food sovereignty is also technological sovereignty, since the affirmation of technological progress in the agricultural sector allows for technological self-determination, as well as acquiring an advantage in terms of food security, thanks to the use of technologies based on distributed registers, i.e. blockchain, which allow for an increase in the level of food safety and supply chain safety and protection of technological and varietal innovations[37].

In relation to the relationship with intellectual property, it is interesting to note the fine line between the usefulness of the law and its overcoming, with the imposition of a lex digitalis. Well, if on the one hand blockchain represents an opportunity for traceability and strengthening the security of the supply chain, on the other hand it could be identifiable as a tool capable of making the rule superfluous, bypassing legal instruments, and moving towards the application also to agriculture, to agritech, of technologies based on blockchain[38] based on the use of algorithms impenetrable by ordinary mortals[39]. This use of technologies is even more alive in the case of data protection, which in themselves are not susceptible to protection based on the discipline of intellectual property, but which can in fact be made inaccessible and not attackable thanks to the introduction of specifically adopted technological barriers[40], such as blockchain which thanks to cryptography and the characteristics that distinguish it allow the protection of data relating to the supply chain and industrial and intellectual inventiveness from cyber attacks.

[36] M. Ferrari, T*he Portici Conference of 2 and 3 December 2022 Food Sovereignty and sustainable innovation between science and law: empathies and dissonances,* in «Rivista di diritto alimentare», (2024), pag. 6.

[37] *Ibid.*

[38] W. D'Avanzo, *Blockchain and smart contracts for the management of the agri-food supply chain. Potential, projects, and legal problems of the Internet of Value,* in «Diritto agroalimentare», (2021) p. 93.

[39] *Ibid.*, p. 10.

[40] G. Spoto, *The uses of Blockchain and the Internet of Things in the food sector*, in «Rivista di diritto alimentare», 2019, p. 25.

In light of the risks represented and in particular the threat to the security of data contained in supply chains and with attention to information relating to inventions, analysts largely confirm that the vulnerabilities existing in the supply chain can be managed by using emerging technologies such as blockchain[41]. In particular, blockchain represents a solution to the gaps in the protection of intellectual property, since its nature as a decentralized and immutable technology radically changes the ways in which property is exercised and the rights connected to it are managed[42].

5. Conclusions

The reflection conducted starts from the identification of two main problems: the need to protect the agricultural and agri-food supply chains from possible prejudicial activities, both of an economic and commercial nature and of a malicious nature; the need to protect the supply chains from cyberattacks aimed to acquiring facts and information concerning the productions and intellectual works related to the agricultural and agri-food sector.

In both cases, the main theme that emerged is the dual nature of blockchain, identified as an optimal tool for protecting the supply chain; in particular, on the one hand, a traceability tool useful for making the supply chain transparent and knowable to third parties, on the other, an immutable and difficult to attack technology thanks to its main feature, which is cryptography, and therefore optimal for protecting the integrity of the supply chain from any possible attempt at misappropriation or malicious attack.

These peculiarities have led not only the doctrine, analysts and operators in the sector to consider it a valid tool to respond to the aforementioned needs, but also by national authorities, such as the National Cybersecurity Agency which identified blockchain in the 2023 Report as a useful technology for strengthening software and platforms[43].

However, the analysis also highlights limitations deriving on the one hand from the risk of a tyranny of technologies over legal rules, but also a necessary implementation of national policies to strengthen new technologies,

[41] C. Mavani, H.K. Mistry, R. Pate, A. Goswami, *The Role of cybersecurity in protecting intellectual property*, cited, p. 25.

[42] *Ibid*., p. 536.

[43] ACN Report 2023, p. 90.

also through investment support policies, if the ultimate goal is really to achieve strategic autonomy in the food and agri-food sector[44].

In any case, in conclusion, the analysis highlights the multiple potential and opportunities of blockchain that simultaneously allows the protection of the national supply chain and the reconstruction of the production process ensuring the authenticity of the product, but at the same time allowing the protection of data and information collected in the supply chain so as to protect against the risks of cyber-attacks, thus satisfying the Cybersecurity needs of the production chains as well as the needs of producers and consumers.

[44] M. Ferrari, *The Portici Conference of 2 and 3 December 2022 Food Sovereignty and Sustainable Innovation between Science and Law: Empathies and Dissonances*, cited, p. 11.

BIBLIOGRAPHY

Agrawal K. , Kumar V.,. Pal R., Wang L., Cheng Y., *Blockchain-based framework for supply chain traceability : a case example of textile and clothing industry* , in «Computers & Industrial Engineering», 2021.

Arisi M., *Cultivating Trust? A Legal Perspective on Data, Blockchain, and Traceability Along the Agri-Food Supply Chain*, in «Quaderni della Rivista di diritto alimentare», (2023), p. 56.

Ballard R., *Blockchain: The facts and Fiction*, in «Oxford Academic Journal», II (2020).

Berke A., *How safe are blockchains? It depends*, in «Harvard Business Review», 2017.

BS Norton, *A guide to Traceability: A practical approaches to Advance Sustainability in Global Supply Chains*, United Nations Global Compact Office, 2014, New York.

Carboni D., *The technologies underlying blockchain, in M. Giordano, R. Battaglini, Blockchain and smart contracts. Operation, legal and international profiles, practical applications,* Milan, Giuffrè, 2019.

Compagnucci L., Lepore D., Spigarelli F., Frontoni E., Baldi M., Di Berardino L., *Uncovering the potential of blockchain in the agri-food supply chain: an interdisciplinary case study*, in «Journal of Engineering and Technology Management», (2022).

Cuccuru P., *Beyond bitcoin: an early overview on smart contracts, in* «International Journal of Law and Information Technology», III (2017).

D'Avanzo W., *Blockchain and smart contracts for the management of the agri-food supply chain. Potential, projects, and legal problems of the Internet of Value*, in «Diritto agroalimentare», (2021).

De Filippi P., Wright A., *Decentralized Blockchain Technology and the Rise of Lex Cryptography,* in «Social Science Research Network», (2015).

De Filippi P., Wright A., Blockchain and the Law Rule of Code, Harvard University Press, Cambridge, 2018.

Ferrari M., *The Portici Conference of 2 and 3 December 2022 Food Sovereignty and sustainable innovation between science and law: empathies and dissonances*, in «Rivista di diritto alimentare», (2024).

Gaetani E., Aniello L., Baldoni R., Lombardi F., Margheri A., Sassone V., *Blockchain-based Database to Ensure Data Integrity in «Cloud Computing Environments»,* in «Proceedings of the First Italian Conference on Cybersecurity» (ITASEC17), (2017).

Iftekhar A., Cui X., Hassan M., Afzal W., *Application of blockchain and internet of things tamper-proof data availability for food safety,* in «Journal of Food Quality», (2020).

Iftekhar A., Cui X., Hassan M., Afzal W., *Blockchain-Based Traceability System That Ensures Food Safety Measures to Protect Consumers Safety and COVID-19 Free Supply Chains*, in «Food», (2021).

Khalkhal B., *IBM Blockchain Transparent Supply & Food Trust. Creating Sustainable Value Chains through Blockchain*, IBM.

Mavani C., Mistry HK, Pate R., Goswami A., *The Role of cybersecurity in protecting intellectual property,* in «International Journal on Recent and Innovation Trend in Computing and Communication», Vol. 12, (2024).

Narayanan A., Bonneau J., Felten E., Miller A., Goldfeder S., *Bitcoin and Cryptocurrency Technologies: A Comprehensive Introduction*, Princeton University Press, 2016.

Natarayan H., Krause S., Gradstein H., *Distributed Ledger Technology (DLT) and Blockchain*, in «World Bank Group FinTech», (2017).

Saberi S., Kouhizadeh M., Sarkis J., Shen L., *Blockchain technology and its relationships to sustainable supply chain management*, in «International Journal of Production Research», (2019).

Spoto G., *The uses of Blockchain and the Internet of Things in the food sector*, in «Rivista di diritto alimentare», (2019).

VALERIA MANZO*

Cybersecurity in law firms between data protection and customisation requirements[§]

Abstract

This paper examines Cybersecurity in law firms in the context of digital transformation and GDPR principles, with a focus on accountability and risk management. Drawing on the IBM Cost of a Data Breach Report 2023, it highlights the legal sector's high exposure to cyber threats. The analysis of common attacks underscores the critical role of human error. The study concludes that tailored security measures and continuous training are essential to effectively protect client data and ensure regulatory compliance.

Il contributo analizza il ruolo della Cybersicurezza negli studi legali alla luce della trasformazione digitale e dei principi del GDPR, con particolare riferimento all'accountability e alla gestione del rischio. Attraverso l'esame dei dati del IBM Cost of a Data Breach Report 2023, emerge l'elevata vulnerabilità del settore legale agli attacchi informatici. L'analisi delle principali minacce evidenzia il ruolo centrale del fattore umano. Ne deriva la necessità di investire in misure di sicurezza personalizzate e in una formazione continua come strumenti essenziali di prevenzione e tutela dei dati.

* PhD at the University of Campania "Luigi Vanvitelli" and Senior Researcher at the Research Centre of European Private Law (ReCEPL).

§ This contribution develops the report presented at the conference 'Cybersecurity and products: challenges and responsibilities in interconnected societies', 29 November 2024, Italian Supreme Court.

1. The use of technology in the practice of law

While the practice of law has always been fraught with critical issues such as legal uncertainty, the difficulty of predicting decision-making processes, the unreasonable length of trials and the effectiveness of judicial protection, many solutions have been proposed with a view to achieving legal certainty, with the judge acting as *an* impersonal *bouche de loi*[1].



Figure 1 – Digital transformation: for the recovery, courage and practicality are needed

With this in mind, and in order not to neglect the efficiency requirements at various levels of the justice system, a series of measures have been introduced to promote the digitisation of court proceedings (to allow remote consultation of files, prevent the risk of tampering, destruction and loss of the documents they contain, and speed up communications between judicial offices and the parties)[2].

Despite the high level of complexity of legal activities and the regulations currently in force, in the current context of digital transformation, they

[1] Below are some fundamental principles enshrined in our Constitution, the European Convention on Human Rights and the Charter of Fundamental Rights of the EU. Art. 3 of the Constitution: 'All citizens have equal social dignity and are equal before the law, without distinction of sex, race, language, religion, political opinion, personal and social conditions'. Art. 111 of the Constitution: "Every trial shall be conducted in the presence of both parties, under conditions of equality, before a third and impartial judge. The law shall ensure that the duration of proceedings is reasonable"; Art. 6 ECHR: "Everyone is entitled to a fair and public hearing within a reasonable time by an independent and impartial tribunal established by law, in order to determine his civil rights and obligations or any criminal charge against him". Art. 21 ECHR "Any discrimination based, in particular, on sex, race, colour, ethnic or social origin, genetic features, language, religion or belief, political or any other opinion, membership of a national minority, property, birth, disability, age or sexual orientation shall be prohibited.". Art. 47 Charter of Fundamental Rights of the EU: "Right to an effective remedy and to a fair trial. Everyone whose rights and freedoms guaranteed by the law of the Union are violated has the right to an effective remedy before a tribunal in compliance with the conditions laid down in this Article. Everyone is entitled to a fair and public hearing within a reasonable time by an independent and impartial.

[2] The same postulate can be inferred from Article 3-*bis* of Law No. 241 of 7 August 1990, which states: 'In order to achieve greater efficiency in their activities, public administrations shall use IT and telematic tools in their internal relations, between different administrations and between these and private individuals'.

prevent replacement of human beings with intelligent agents[3], it is legitimate to note that the appropriate use of technology by legal practitioners can be a determining factor in reducing some of the critical issues identified.

It is therefore necessary to consider technology not as a substitute for legal professionals but as a resource/tool.



Figure 2 – Artificial Intelligence as a resource or as a fear? | Centro di Orientamento Pastorale

On this point, authoritative doctrine[4] has carefully observed that intelligent machines should be understood in the same way as simple technological tools designed to replace human tasks that which, presumably, will continue to be performed by lawyers.

In such a context, the questions posed by Lawlor are particularly relevant and forward-looking.

In regard to the impact of new technologies on society[5], asked: *"Will computers revolutionise the practice of law and the administration of justice, as they will in almost everything else? Will they help the legal profession in the analysis of legal materials? Will they help make law less unpredictable?".*

[3] L. Viola, *Giurimetria, composizione di dati e giustizia predittiva*, in «La Nuova Procedura Civile», III (2019).

[4] P. Moro, *Intelligenza artificiale e professioni legali. La questione del metodo,* in «Journal of Ethics and Legal Technologies (JELT)», vol. 1, I (2019), pp. 24-43.

[5] C. Lawlor Reed, *What Computers Can Do: Analysis and Prediction of Judicial Decisions,* in «American Bar Association Journal», vol. 49, IV (1963), pp. 337-344.



Figure 3 – Digital Justice in the Era of Covid-19: a Pan-European Perspective – 24th June 2020, 12.00pm BST | Cooke, Young & Keidan

In the field of digital justice, a study on the digitisation of justice conducted by Bucherius Law School and the Boston Consulting Group[6], concluded that the following three tools are necessary: enabler technologies, which include the cloud and the software needed to ensure basic levels of Cybersecurity; support process solutions such as tools, mainly management tools (case management and document and knowledge management), aimed at optimising the management of procedural activities such as the creation and consultation of digital files; and, lastly, substantive law solutions, which are the most advanced level of legal tech tools capable of supporting legal practitioners in carrying out their traditional tasks of interpreting and applying substantive law[7].

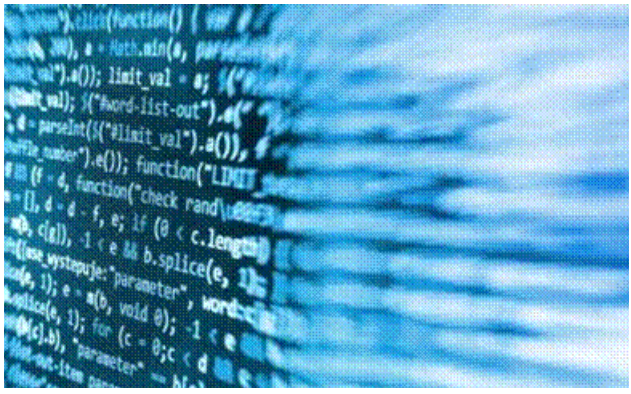


Figure 4 – Algorithms e augmented justice | Università Cattolica del Sacro Cuore

[6] D. Hartung *et al.*, *The future of digital justice*, in «Bucerius Law School Studies», (2022).

[7] This category includes *document automation tools* such as advanced legal research *software* and predictive justice.

However, from the perspective described above, we note the consolidation of various realities heterogeneous legal tech tools that contribute to defining the scope of so-called 'augmented justice'.

2. The principles of the GDPR

EU Regulation 2016/679 on the protection of personal data not only reaffirms the fundamental principles already in force under previous legislation but also adds new ones.



Figure 5 – Privacy Regulation, 10 things you need to know about GDPR

Among the principles relating to data processing that are confirmed are:

– purpose of processing, which limits its use solely to the purposes of protection, advice and defence pursued with a specific mandate from the solicitor (data controller)[8];

– necessity and proportionality: processing must be adequate, relevant and necessary for the purpose[9];

– limited duration: processing may not continue beyond the time necessary to carry out the tasks, i.e. beyond the time necessary for the mandate and the purpose of the processing itself, including legal storage obligations;

– respect for the rights of individuals;

– security and confidentiality: the solicitor is required, also due to ethi-

[8] For example, data collected in land registry searches cannot be used to learn about the private lives of clients, nor for commercial, political or electoral advertising purposes.

[9] Case files and their electronic archiving, for example, must be configured in such a way as to minimise the use of personal and identifying data.

cal obligations and in compliance with professional secrecy, to provide an adequate level of security for the data of their clients, taking all necessary measures to ensure the confidentiality, integrity and availability of personal data[10].



Figure 6 – GDPR and accountability

Additional principles and duties have also been introduced with which lawyers must comply, such as the principle of accountability, data minimisation, the right to be forgotten, impact assessment and the right to data portability.

A milestone in a different approach to data subjects' data is represented by the principle of *accountability,* which requires (among others) lawyers to undergo a profound cultural change in the way they process information that comes into their possession or to which they have access by virtue of their mandate.

With regard to the Privacy Code, the so-called minimum measures are no longer required, but the data controller is responsible for defining, after careful analysis of the data processed and the possible risks involved, adequate measures to ensure compliance with the provisions of the GDPR.

Essentially, accountability means that measures must be tailored to the structure of the individual controller and developed on a case-by-case basis, using prior mapping of the data processed and the risks involved in data processing.

But *accountability* also means being able to 'give an account' of the activities carried out, of compliance with the principles of the GDPR and of *compliance* with the Regulation on the processing carried out.

[10] The data contained in the file, for example, cannot be consulted by persons who are not qualified and expressly instructed and authorised to access it because of their specific tasks, whether they are internal to the law firm (secretarial staff, trainees, colleagues) or external to it (co-defence lawyers, technical consultants, accountants).

Data minimisation is the principle that the personal data to be processed for each individual activity must be only those necessary to achieve the purpose.

In order to comply with this principle, solicitors must, as far as possible, process only data that is considered essential, necessary and relevant to perform the service requested by the client.

Article 17 of the GDPR provides for the right of the data subject to obtain the erasure of personal data concerning him or her from the data controller's files and the corresponding obligation to comply without undue delay (the so-called right to be forgotten).

Article 35 of the GDPR requires that, where a type of processing is likely to result in a high risk to the rights and freedoms of natural persons, the data controller must carry out a prior impact assessment (DPIA).

Therefore, the more criteria the processing meets, the more likely it is that there is a significant and high risk to the rights and freedoms of data subjects and that it is therefore necessary to carry out a data protection impact assessment.

Finally, the right to portability gives data subjects the right to request that the data controller transfer their data to another controller, without the former being able to object.

Specifically, Article 20 of the GDPR gives data subjects the right to receive, in a structured format, the personal data concerning them that they have provided to the data controller, as well as to transmit those data to another data controller without hindrance from the first data controller when the processing is based on consent or on a contract and is carried out by automated means (e.g. using computerised management software); all in compliance with the provisions of Article 2235 of the Italian Civil Code[11].

3. The IBM 2023 Security Report

The recent IBM Security report eleased the Cost of a Data Breach Report, highlighting how the global average cost of Data breaches reched $4.45 million in 2023 – an all-time high for the report – with detection costs up 42% compared to the same period last year.

[11] So-called prohibition of retention: "*The service provider may not retain the items and documents received, except for the period strictly necessary to protect their rights in accordance with professional laws*".

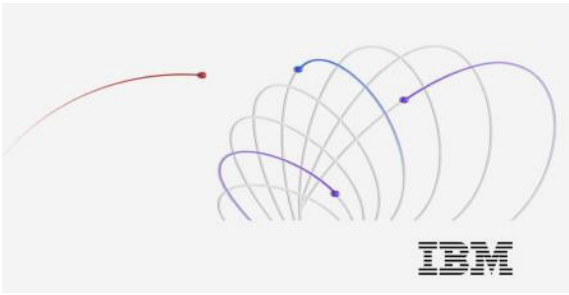


Figure 7 – Cost of a Data Breach: IBM report about the escalation of data violation

The financial in the professional services sector, including the legal sector, has therefore, revealed a disarming fact.

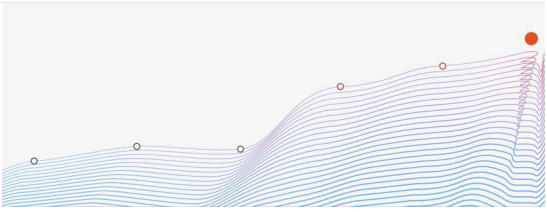


Figure 8

As stated by Chris McCurdy, General Manager, Worldwide IBM Security Services: *"Time is the new currency in Cybersecurity. As the report indicates, early detection and rapid response can significantly reduce the impact of a breach", (...) "Investments in threat detection and rapid response, thanks to AI and automation, are essential to mitigate attacks as effectively as possible".*



Figure 9 – Map of Italy: political and physical differences – UniD Formazione

At the Italian level, the 2023 Report revealed interesting insights into the country's situation: 1) the average total cost of data breaches is €3.55 million[12]; 2) on average, it takes 235 days to identify and contain a cyber threat[13]; 3) the main attack vectors are: social engineering, phishing and stolen or compromised credentials; 4) Artificial Intelligence and automation had the greatest impact on the speed of identification and containment of breaches in the companies and law firms surveyed; and 5) nearly 41% of the data breaches analysed resulted in data loss in multiple environments, including public cloud, private cloud and on-premise environments, demonstrating how cybercriminals were able to compromise multiple environments while avoiding detection[14].



Figure 10 – Lawyer exam is now a chaos. The protest of the practitioners – Corriere.it

This raises the question of why the legal sector is a prime target for cybercrime.

While it is true that law firms are custodians of a wide range of highly confidential and sensitive information, understanding the common types of cyber attacks that threaten the sector is a key step in strengthening Cyber security.

4. So-called 'common' attacks in the legal sector and the importance of Cybersecurity training

The threats that can affect law firms can be divided into external threats and those resulting from human error.

[12] An increase compared to €3.40 million in 2022.

[13] This is 15 days less than the Italian average of 250 days in 2022.

[14] It goes without saying that data breaches affecting multiple environments also led to higher breach costs (€3.72 million on average).

The former include, *among others, ransomware* and USB attacks; the latter include *phishing*, BEC attacks, *vishing* and *smishing*.



Figure 11 – Mengenal Apa Itu Ransomware, Cara Kerja, dan Jenis-Jenisnya-Tekno Tempo.co

Let us analyse the main ones.

Ransomware (a type of malware) is a particularly insidious threat to the legal sector and extremely profitable for those who use it, as it can encrypt and steal data from the entire firm's network stored on PCs, rendering it inaccessible.

This is confirmed by the Clusit 2024 Report, which found that in 81% of cases, the severity of the attacks is, according to the severity scale used, and that *ransomware* accounts for two-thirds of cyber attacks, resulting in a significantly worse picture than in previous years[15].

Specifically, once cybercriminals have stolen the data, they demand a ransom for the decryption key or for not publishing the sensitive data they have obtained *online.*

What protective measures can therefore be taken? The best protection is prevention, i.e. always updating both the antivirus and the operating system, implementing advanced protection and detection systems (IDS, IPS, EDR).

This is followed by *the backup* of the firm's data, planned according to *security by design*[16].

[15] https://clusit.it/rapporto-clusit/.

[16] *Backups* should always include 'redundancy', meaning not just one *backup* copy but at least three copies according to the basic 3-2-1 rule, i.e. three copies of each piece of data you want to keep, two of which are 'onsite' but on different storage devices (HD, NAS, Cloud, etc.) and one 'off-site' copy (at a remote location) on the Cloud. This way, if *ransomware* infects the system, a copy of the data will remain protected, allowing for recovery if necessary.



Figure 12 – Usb attack which damage usb

The same cannot be said for USB attacks, which can manifest themselves in the following ways: 1) reprogrammable microcontrollers (where the internal microcontroller of the USB device is reprogrammed and automatically performs keyboard operations); 2) *firmware attacks* (where the *firmware software i*nside a USB device is programmed to perform harmful actions such as automatically *downloading* infected files) and 3) unprogrammed USB attacks (in which case the USB device may appear empty, but there are hidden files inside that will be used to secretly transport confidential data).

So how can USB attacks be prevented? First of all, by creating secure USB keys using external *software* that allows files to be encrypted internally, or by using qualified antivirus software, or by disabling the auto-run function on all devices.



Figure 13 – Phishing ad its variations: here are the countermeasures to adopt – Cyber Security 360

Moving on to threats arising from human error, phishing attacks involve cybercriminals using e-mails, text messages or hyperlinks to lure victims to malicious websites that download malware onto their computers.

Phishing attacks involve cybercriminals using emails, text messages or hyperlinks to lure victims into visiting malicious websites that can download *malware* onto their computers or steal personal information such as login details.

Specifically, in the legal sector, it is clear how such attacks can compromise sensitive case data and client information.

Here too, what are the remedies? It is of primary importance to check the website link (URL) as it is very rare for the https protocol (which encrypts data) to be used. it is also necessary to check the DNS (which allows you to connect to websites using domain names instead of IP addresses) and, finally, to check short (unreadable) links using various web tools that allow you to 'unwind' the link to show its actual origin and verify that it is not a malicious link.



Figure 14 – Business Email Compromise BEC: how to prevent and to protect oneself

BEC (Business Email Compromise) attacks are a 'sophisticated form of phishing' whereby cybercriminals attempt to trick executives or budget holders into transferring funds or revealing sensitive information. Law firms that handle large financial transactions are the primary targets of such attacks.

With the increasing digitisation of tax and institutional services, *online* scams and attempts to steal access to current accounts are becoming increasingly common.

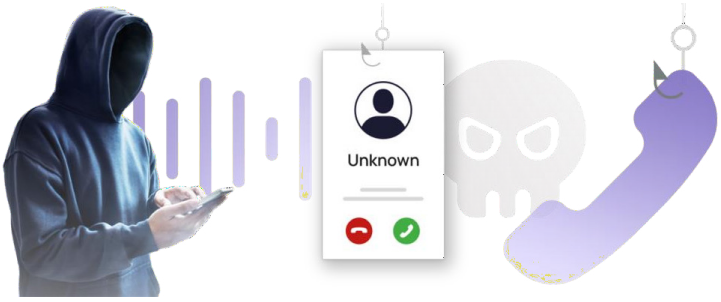


Figure 15 – Vishing: suggestions of the Guarantor to protect oneself from mobile Phishing – AssoPrivacy

Among these is *vishing* (a term derived from the combination of 'voice' and 'phishing') has become a common technique for fraud that aims to obtain sensitive information from users through fake phone calls.

Specifically, fraudsters simulate calls from financial institutions, persuading recipients to act against their own interests and provide access data to current accounts and other personal details, thus jeopardising the security of online transactions.

If you suspect you have been the victim of such a scam, you must take appropriate measures to protect your identity and login details.

First of all, hang up the phone and avoid providing or confirming any personal or financial information. Then report the fraud to your bank, providing all available details about the suspicious call.

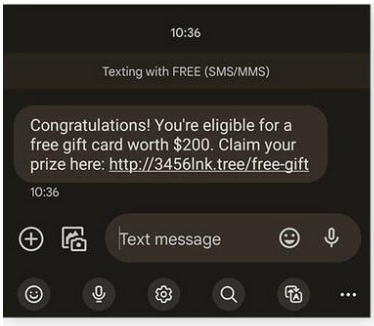


Figure 16 – What is smishing? Protect yourself from SMS phishing

Smishing is one of the various forms of so-called 'social engineering', i.e. the methods used by cybercriminals to steal sensitive information, often for financial gain, through text messages (including instant messages).

Our Data Protection Authority provides guidance on its website on how to defend yourself against such attacks: first and foremost, the Authority points out that neither the public administration nor banks and insurance companies ever ask their customers for confidential information by e-mail, SMS or telephone, emails, text messages or telephone. Without prejudice to this general warning, it suggests never communicating personal data and information to strangers; not storing bank or credit card credentials on smartphones; frequently checking transactions and, if necessary, activating automatic *alert systems* that notify users of every transaction carried out; do not click on *links* contained in messages or open any attachments and, lastly, check whether the text contains any anomalies, such as linguistic, grammatical or lexical errors, etc. (often signs of *smishing).*

Given the above, in a scenario such as the one outlined above, where the weak link in the chain is the human factor, investing in training in the legal sector to raise concrete awareness of security is a key pillar of defence against increasingly frequent cyber threats.



Figure 17 – TRAINING, WHAT DO YOU NEED? (download the form for your training needs) – Confcommercio Foggia

In addition to providing the necessary skills to effectively recognise and mitigate these vulnerabilities, it also ensures that those working in a law firm become aware of their obligations under data protection regulations, guaranteeing their clients the utmost care in the handling of confidential information.

Security awareness training is a fundamental component of any Cybersecurity strategy, aiming to create a 'culture of security within the law firm'.

How can effective training be implemented?

As cyber threats are constantly evolving, *one-off* training is not sufficient; regular and up-to-date training sessions are required.

It should also be remembered that each member of the law firm has unique responsibilities and access to different types of data; for this reason, it is necessary to tailor training to the specific needs of the different roles involved[17].



Figure 18 – Courses and trainings on Misura | Ufficio Service

[17] For example, partners and support staff may require separate training modules that address their respective roles, responsibilities, and potential cybersecurity risks.

Moving forward, implementing interactive elements in training is an activity that can actively engage participants and offer practical experience in dealing with potential cyber threats; this is a proactive way to strengthen preparedness.

5. Final considerations

If, as has already been correctly recognised[18], the advent of digital technology *is changing the way we see ourselves,* altering human habits with significant implications, including legal ones[19], then perhaps the risk we face is that of preserving an anachronistic right in the face of technological evolution.

Private law itself, through a decisive interaction between contract law, personal data protection, digital identity and intellectual property, is increasingly called upon to face complex challenges that continue to challenge its traditional balance[20].



Figure 19 – Digitalization and legal office: challenges and opportunities | TeamSystem

It however, try to reflect on possible future prospects.

Cybersecurity in the legal sector is not an option but a necessity.

Protecting client data, maintaining trust and ensuring compliance with data protection regulations. Data has become more fundamental than ever.

The objective is therefore to contain risk, but it seems clear that it cannot be eliminated entirely, especially in the specific area we are dealing with.

A possible scenario could be to provide for a form of *co-regulation* in which the rules are suggested by the market (perhaps through the contribu-

[18] L. Florindi, *Il verde e il blu. Idee ingenue per migliorare la politica,* Raffaello Cortina Editore, Milano 2020.

[19] M. Cartabia., *"Nuovi diritti" e leggi imperfette*, in «Iustitia», (2016), p. 170.

[20] F. Bocchini, E. Quadri., *Diritto privato, 7ª* ed., Giappichelli, Torino 2018, p. 48.

tion of *stakeholders* in the sector and scholars in the field) and shared by the sector authorities, following the "regulatory circle' model[21].



Figure 20 – Best-practice – AICQ Centronord

Given the rapid evolution of the cyber threat landscape, law firms must therefore adopt *best practices* aimed at remaining vigilant and safeguarding their digital resources and protecting their clients' sensitive information.

[21] E. Del Prato, Qualificazione degli interessi e criteri di valutazione dell'attività privata funzionale tra libertà e discrezionalità, in «Rivista trimestrale di diritto e procedura civile», (2012), p. 411 ss.

BIBLIOGRAPHY

Bocchini F., Quadri E., *Diritto privato*, 7ª ed., Giappichelli, Torino 2018, p. 48.

Cartabia M., *"Nuovi diritti" e leggi imperfette*, in «Iustitia», II (2016), p. 170.

Del Prato E., *Qualificazione degli interessi e criteri di valutazione dell'attività privata funzionale tra libertà e discrezionalità*, in «Rivista trimestrale di diritto e procedura civile», (2012), p. 411 ss.

Floridi L., *Il verde e il blu. Idee ingenue per migliorare la politica*, Raffaello Cortina Editore, Milano 2020.

Hartung D. et al., *The future of digital justice*, in «Bucerius Law School Studies», (2022).

Lawlor Reed C., *What Computers Can Do: Analysis and Prediction of Judicial Decisions*, in «American Bar Association Journal», vol. 49, IV (1963), pp. 337-344.

Moro P., *Intelligenza artificiale e professioni legali. La questione del metodo*, in «Journal of Ethics and Legal Technologies (JELT)», vol. 1, I (2019), pp. 24-43.

Viola L., *Giurimetria, composizione di dati e giustizia predittiva*, in «La Nuova Procedura Civile», III (2019).

FEDERICA MARANI*

The relationship between GDPR and blockchain: between data privacy and technological security§

Abstract

The advent of the Internet has led to the emergence of new issues that have required and still require a delicate balancing of opposing interests. Today, it is possible to harness a plurality of new technologies. However, for such tools to be efficiently employed and developed, it is necessary to understand whether and how they can be grafted within the legal system, in order to avoid a disruption of the "traditional" categories and "general" disciplines, as their compliance is fundamental to guaranteeing the rights and guarantees of individuals confronted with such new technologies, including from the perspective of data protection. From both legal and technological perspectives, therefore, it is necessary to develop tools that ensure that individuals have effective control over their rights and, specifically, over their personal data.

L'avvento di Internet ha comportato l'insorgere di nuove problematiche che hanno necessitato e tuttora necessitano un delicato lavoro di contemperamento tra interessi contrapposti. Oggi, ci si può giovare di una pluralità di nuove tecnologie che offrono potenzialità impressionanti. Tuttavia, affinché tali strumenti possano essere efficientemente impiegati e sviluppati, è necessario comprendere se e come essi si possano innestare all'interno dell'ordinamento per evitare un disgregamento delle categorie "tradizionali" e delle discipline "generali", il cui rispetto è fondamentale per tutelare i diritti e le garanzie degli individui che si confronta-

* PhD student in "Diritto e management della sostenibilità" at Siena University.

§ This contribution develops the presentation at the conference 'Cybersecurity e prodotti: sfide e responsabilità nelle società interconnesse',29 November 2024, Italian Supreme Court.

no con tali nuove tecnologie, anche sotto il profilo della protezione dei dati personali. Tanto sul versante legale quanto su quello tecnologico, dunque, è necessario elaborare strumenti volti ad assicurare un effettivo controllo dei soggetti sui propri diritti e, nello specifico per quanto qui in esame, sui propri dati personali.

1. Introduction. The processing of personal data

The focus on personal data's protection is becoming increasingly more significant, catalyzing the interest of almost all legal systems, at national, supranational, and international level.

The increasingly fast and immediate exchange of information and personal data on the web has necessitated the introduction of new regulatory mechanisms, aimed at balancing opposing interests that are in constant dialectic with each other: on the one hand, the analysis and use of personal data, which is necessary for technological development and economic growth, and on the other hand, respect for the fundamental rights of citizens and, in particular, the guarantees of individuals operating in the new market: the so-called 'digital consumers', who must be protected, perhaps even more than in the 'traditional' (off-line) context, and considered as weak subjects in the relationship with providers of digital services. than in the 'traditional' (off-line) context and considered as weak subjects in the exchange relationship with digital service providers.

The entry into force of the Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016, the 'General Data Protection Regulation' (so-called GDPR) stands as the culmination of the story that began, in Europe, in 1995, which led to the introduction of a uniform and generalized protection of the right to privacy and the movement of personal data, directly applicable in all Member States.

The aforementioned Regulation hinges on a number of basic principles: on the one hand, with the aim of increasing the accountability of the parties involved in the processing of personal data (data controller and data processor), the principles of accountability, data protection by design and by default, pseudonymisation and minimisation have been introduced and intensified, which have given the data controller a new active role as guarantor of compliance with the principles laid down by the regulation, through the inclusion of complex duties of management and prevention of the specific risk associated with the processing of personal

data[1]. With the introduction of the principle of accountability, in essence, the very concept of 'risk' has changed, which is no longer limited to the breach phase alone, but also extends to the phase preceding the processing itself. On the other hand, in order to provide those concerned with greater awareness in the use of their personal data, the GDPR consolidates the fundamental principles on the protection and processing of personal data, some of which were already present in the 'old' Privacy Code, while others have been specifically introduced (i.e. the principle of lawfulness of processing, the principle of transparency, the principle of accuracy, the principle of limitation of storage, the principle of information, the principle of consent, the principle of access, rectification, limitation and objection, the principle of security and integrity, the principle of erasure, and the principle of portability).

Significantly, among others, a new liability mechanism for unlawful processing of personal data was introduced in Article 82 of the Regulation, which grants compensation protection to anyone who suffers material or immaterial damage caused by a breach of the Regulation on the part of the data controller or processor, unless the latter provides evidence that the damaging event is in no way attributable to them.

Having briefly outlined the current legal framework on the processing of personal data, we propose to investigate and verify whether this state of the art, based on the remedies introduced by the national and supranational legislator, is at least adequate to ensure effective protection for those operating on the digital market and interfacing with new technologies, particularly in the field of smart contracts operating on blockchain systems.

2. The "new technologies": smart contracts and blockchain

The definition of smart contracts was provided in 1994, by the North American computer engineer, cryptographer, and legal expert Nick Szabo, who understood them as computer protocols aimed at facilitating, verifying and

[1] Cf. E. Tosi, *Responsabilità civile per illecito trattamento dei dati personali e danno non patrimoniale. Oggettivizzazione del rischio e riemersione del danno morale con funzione deterrente-sanzionatoria alla luce dell'art. 82 GDPR*, Giuffré, Milano 2019, p. 40, where the A. pointed out that, with the principle of accountability, the passage «from *ex post* responsibility alone to conscious and documented *ex ante* accountability» has been marked, as this principle «si estrinseca nella valutazione, gestione e prevenzione del rischio *ex ante* da parte degli attori del trattamento – titolare e responsabile nei casi previsti dal GDPR – controbilanciato da un controllo *ex post* da parte dell'Autorità di protezione dei dati e giudiziaria: quest'ultima competente per l'accertamento e liquidazione del danno – patrimoniale e non patrimoniale – ai diritti fondamentali della riservatezza, della protezione dei dati e dell'identità derivante da illecito trattamento dei dati personali».

enabling automated negotiation in which the conclusion and execution of the contract takes place by means of so-called self-executing clauses, also activating automated remedies in the event of non-performance[2].

Smart contracts operate through distributed ledgers (so-called Distributed Ledger Technology-DTL[3]); specifically, in the blockchain[4].

The blockchain is a technology based on a shared public register, within which a series of data protected by cryptographic keys are entered by interconnected devices (so-called system-nodes), which are stored and organized in a chronologically ordered chain of 'blocks', identified by means of an alphanumeric string called a 'hash', which also includes the hash of the previous block, to form an unchangeable and unalterable concatenation.

The characteristics of this technology, designed to guarantee the security and immutability of information, make any attempts to tamper with the information it contains[5] highly costly. The blockchain is, in fact, a so-called 'append-only' database: once a piece of data is entered, it cannot be deleted and there is no way to go back to the previous step[6]. All this, in a decentralized system: that is, negotiation operations take place automatically and without the intervention of any third-party intermediary[7].

When words are recorded on the blockchain, they are transformed into codes: language is transformed into in-put through the use of so-called

[2] Cf. T. Pellegrini, *Gli smart contract, in Diritto privato digitale,* a cura di E. Battelli, Giappichelli, Torino 2024, pp. 359 ff., where the author points out that the first definition of smart contract shows that the aim of the instrument was to reduce reliance on the spontaneous cooperation of the contractual counterparty: if the contract executes itself thanks to a computerized protocol, Titius will receive what Caius owes him regardless of the latter's spontaneous action.

[3] M. Maugeri, Smart Contracts e disciplina dei contratti, Il Mulino, Bologna 2021, pp. 28 ff.: the acronym DLT refers to a technology that enables the recording and storage of data through multiple repositories (ledgers), each of which simultaneously contains the same data that are stored and controlled by a network (nodes).

[4] D. Macrinici, C. Cartofeanu, S. Gao, *Smart contract applications within blockchain technology: A systematic mapping study,* in «Telematics and Informatics», VIII (2018), pp. 2337-2354. Regarding smart contracts, the authors states that «*they are just like contracts in the real world. The only difference is that they are completely digital, which means a small programing code that is stored inside of a blockchain*».

[5] N. Travia, *La tecnologia blockchain, in Diritto privato digitale,* a cura di E. Battelli, cit., pp. 338 ff.

[6] K. Yeung, *Regulation by Blockchain: The Emerging Battle for Supremacy between the Code of Law and Code as law*, in «The Modern Law Review», II (2019), pp. 207 ff.

[7] Therefore, as observed by M. Proto, *Questioni in tema di intelligenza artificiale e disciplina del contratto, in Il diritto nell'era digitale. Persona, mercato, amministrazione, giustizia*, a cura di R. Giordano, A. Panzarola, A. Police, S. Preziosi, M. Proto (eds.), Giuffré, Milano 2022, p. 181, it deals with «database – anche detti nodi, decentralizzati e consultabili da utenti – strutturati in blocchi, nei quali possono essere inseriti valori digitali o rappresentanti beni materiali, in modo che sia facile identificare l'attuale proprietà e registrarne i trasferimenti, al pari di un registro notarile».

'computer syntax', which is based on the use of programs aimed at transforming 'natural language' into 'machine language'.

In summary, the pillars of the blockchain architecture are, on the one hand, the distributed ledger and, on the other hand, cryptography. Alongside these, however, stands a third one: consent, to be understood as adherence to the distributed ledger by users who, once adhered to, will operate in the "virtual world" of the blockchain and respect the rules underlying this technology[8]. This is the reason why, in a broad sense, control by an authority is not necessary: the system legitimizes itself through sharing between participants[9].

The smart contract is, therefore, a type of contract written with an algorithmic language (one speaks in this regard of a cryptographic lex[10]) and stored on a blockchain[11]. The code stored on the blockchain describes, therefore, the contractual terms, including the good or service that is the subject of the contract. The system is based on the so-called 'if this, then that' logic: if the envisaged in-put is realized, then the out-put will follow[12]; when an ascertainable condition occurs (if this...), the digitally ascertainable effect is produced (then that...)[13].

Once a smart contract is created and stored on the blockchain, the code of the contract is made public and viewable by anyone; the data associated with the contract, on the other hand, is encrypted and can only be consulted by those who have private keys (this is referred to as "asymmetric encryption"[14]).

From a regulatory point of view, the first Italian intervention on the matter was recorded with Decree-Law No. 135 of 14 December 2018 (so-called 'Simplification Decree', converted with amendments into Law No. 12 of 11

[8] On this point, cf. E. Gabrielli, *La nozione di contratto e la sua funzione. Appunti sulla prospettiva di una nuova definizione di contratto*, in «Giust. civ.», 2019, 1, pp. 299 ff., according to which in smart contracts operating on blockchain systems, the agreement between the parties is replaced by «mere unilateral declarations» inserted within a system of algorithms and cryptographic codes and, as a result, the contract is structured by means of algorithms representing the transposition in computer terms of the will of the parties to conclude a given contract and to subsequently execute it.

[9] M. Callegari, *Sostenibilità, supply chain e intelligenza artificiale,* in «Giur. it.», V (2024), p. 1217.

[10] F. Longobucco, *Utopia di un'autonoma Lex Criptograp(h)ica e responsabilità del giurista,* ESI, Napoli 2023.

[11] Cf. T. Pellegrini, *Gli smart contract,* cit., p. 364, who points out that, in the current context, *only what can enter a blockchain can be the subject of smart contracts.*

[12] B. Sirgiovanni, *Lo "smart contract" e la tutela del consumatore: la traduzione del linguaggio naturale in linguaggio informatico attraverso il legal design,* in «Nuove Leggi Civ. Comm.», I (2023), pp. 215 ff.

[13] G. Bevivino, *Smart contract e "assestamenti" della disciplina generale del contratto,* in «Contratto e impresa», II (2024), p. 393; C. Attanasio, Inadempimento dello smart contract, sistema rimediale e tutela effettiva, in «Riv. dir. civile», IV (2024), p. 719 ss.

[14] As pointed out by A. M. Gambino, C. Bomprezzi, *Blockchain e protezione dei dati personali*, in «Dir. Inf.», III (2019), p. 626, the public key coincides with the user's account and, as public, can be

February 2019) which, in Art. 8-ter, paragraph 2, defined smart contract as a computer program that operates on distributed ledger-based technologies and the execution of which automatically binds two or more parties on the basis of effects predefined by them and, in paragraph 1, offers a definition of «distributed ledger-based technologies» – among which blockchain is included –, in the part where it provides that, by this locution, it refers to information technologies and protocols that use a shared, distributed, replicable, simultaneously accessible, architecturally decentralized ledger on a cryptographic basis, such as to enable the recording, validation, updating and storage of both unencrypted and further cryptographically protected data that is verifiable by each participant, not alterable and not modifiable.

2.1. Blockchain permissionless and permissioned, public and private: differences and implications

Characteristic of the blockchain operating system is, therefore, the sharing of rules common to all participants. These rules, it should be noted, are not the same in all cases: they differ according to the heterogeneity of the distributed registers, which may be private or public, centralized or decentralized, subject or not subject to authorization or authority[15].

The network, first of all, is characterized on the basis of the pre-established access mode. It is customary to distinguish between permissionless and permissioned blockchain.

The permissionless blockchain is usually used to manage transactions in cryptocurrencies (e.g. Bitcoin, Ethereum, etc.). In such a system, anyone can join the network, and any computer can act as a node, without needing to be authorized; the network can thus grow by leaps and bounds, and, given the very high number of nodes, the consensus mechanism – which would require the agreement of 50% plus one of the nodes, which tends to be impossible – makes this type of blockchain truly immutable. Consequently, the network is completely transparent with respect to the information circulating within it, as any user can view the entire blockchain[16].

In permissioned blockchains, on the other hand, there is the presence of centralized models of network access control, based on protocols for validating transactions (so-called pre-identification) carried out

known by everyone. It is then matched with a 'private' key, which as such remains secret. In this way, messages intended for a specific person are encrypted with his public key, so that he can only access them with his private key.

[15] Cf. C. Poncibò, *Il diritto comparato e la blockchain*, ESI, Napoli 2020.

[16] A.M. Gambino, C. Bomprezzi, *Blockchain e protezione dei dati personali,* cit., p. 627.

by a predefined set of 'trusted'[17] subjects. Access control means that the identities behind the nodes are known, and this decreases the level of immutability, as it is abstractly conceivable that the majority required to make changes is reached. For these reasons, such typologies can be profitably used in the business service sector and public administration. Therefore, transparency is also diminished[18].

To summarize, the main differences between the two models lie in the identifiability of the actors using them, the way nodes are selected and the number of users, the so-called shared consensus mechanism and the transparency of the block content[19].

Another important difference, which is based on the management of the IT infrastructure, is that between public and private blockchains.

Public blockchains are characterized by the lack of any central control, and the exchange of information and data is governed by the participants themselves through the expression of digital consent. Thus, public blockchains are self-governed, and the network is created as participants are added. Private blockchains, in contrast, are managed by a single organization or group of individuals. Normally, public blockchains are also permissionless, but this does not detract from the fact that they can also be permissioned if they contain pre-identifications on entry; for private ones, on the other hand, it is extremely difficult to imagine a permissionless system.

This being said, the purpose of the blockchain is to disintermediate exchanges[20], and this is possible precisely because of the features outlined above: the use of smart contracts, to manage the exchanges of information between nodes; the absence of a central authority, which means that exchanges are entirely managed by the network participants by means of special computer protocols for the expression of consent and the validation of the transactions carried out; a high level of security, ensured (rectius, hopefully ensured) through encryption and the tendency for the distributed ledger to be unchangeable[21].

[17] P. Rubechini, *Criptoattività e blockchain*, in «Giornale di dir. amm.», II (2023), p. 263-264.

[18] A.M. Gambino, C. Bomprezzi, *op. cit.*, p. 627.

[19] *Ibidem.*

[20] V. N. Travia, *La tecnologia blockchain, cit., p. 342,* according to which «le caratteristiche della blockchain sono considerate ideali (in termini di qualità, sicurezza, integrità, immodificabilità) per realizzare l'obiettivo che la maggior parte degli ordinamenti persegue mediante le discipline relative alle sottoscrizioni e ai documenti elettronici: quello di conferire rilevanza giuridica alle rappresentazioni informatiche e di determinare la specifica efficacia giuridica agli effetti sostanziali e probatori, nei rapporti con le pubbliche amministrazioni e tra privati».

[21] P. Rubechini, *Criptoattività e blockchain*, cit., pp. 263-264.

3. Is it possible to find a connection between GDPR and blockchain technology?

The scope of the blockchain is extremely broad: potentially, any data can be recorded within it[22].

In this regard, it should be recalled that there is no unanimity of views with reference to the very qualification of data entered in the blockchain. According to some, it would be "anonymous" data, as such removed from the application of the GDPR, since the public keys appear encrypted, so the identity of the subject remains hidden. However, if this is true, according to other doctrine[23] it would more correctly be pseudo-anonymous data, as there are techniques for tracing the address of the key and, once the public key address is linked back to an identity, it is possible to trace the transactions involving that identity.

Although the debate is open, it seems reasonable to adhere to this thesis and identify the data in question as "pseudo-anonyms," subject to the GDPR and identified as information about an identifiable natural person, since it can be traced back to that person through the use of "further information" (Art. 4(1)(5) GDPR).

Assuming that processing of personal data on blockchain applications (so-called "DApps") takes place, it is necessary to understand what the implications and relationship between this technology and the European Data Protection Regulation are.

At first analysis, the regulations under the GDPR seem to be incompatible with the decentralized and immutable system of the Blockchain.

However, it is precisely the decentralized and immutable structure that could ensure the security and confidentiality of personal and sensitive information, reducing the risk of data manipulation and ensuring, through the cryptographic system, high data privacy and security.

In particular, the characteristics of blockchain are found to meet the requirements of the GDPR about compliance with the principle of Privacy by Design and Accountability.

Article 25 GDPR, concerning data protection by design, requires the data controller to put in place "appropriate technical and organizational mea-

[22] As has been effectively observed by A.M. Gambino, C. Bompezzi, op. cit., p. 624, the functionalities of blockchain are innumerable. Not only money, but potentially any data can be recorded, as well as personal data. The European Parliament, back in February 2017, published a report highlighting the application areas in which blockchain could have greater use. These includes virtual currencies, patents, e-voting, smart contracts, product supply chains, public services.

[23] A.M. Gambino, C. Bompezzi, op. cit., pp. 633 ff.; M. Giuliano, *La blockchain e gli smart contracts nell'innovazione del diritto nel terzo millennio*, in «Dir. Inf.», VI (2018), pp. 989 ff.

sures" to ensure compliance with the principles enshrined in the Regulation. In summary, measures are preferred that ensure minimization of processing, use pseudonymization techniques, ensure transparency in the use of data, allow the data subject to exercise effective control over his or her data, and ensure a high level of security.

From this perspective, the characteristics of blockchain are suitable for meeting these requirements[24]. In fact, this technology – especially in its "public" form – would make it possible to achieve, through the system of so-called hashes, an almost total security and integrity of the collected data. In addition, through the mechanism of asymmetric encryption, selective access to data is guaranteed, respecting what users have actually allowed due to the purpose of processing, in line with compliance with the principle of minimization.

The characteristics of traceability and immodifiability, moreover, enable data subjects to be able to monitor access to their data, ensuring maximum recognition of the principle of transparency. The combination of these prerogatives would facilitate the data controller himself to comply with the principle of accountability.

Nonetheless, the countless friction points with the GDPR Regulation itself cannot be overlooked, especially about the structure of permissionless and public blockchains, to which this analysis is limited[25].

First of all, the fact that information, albeit encrypted, is made known to the public on the blockchain if, on the one hand, it guarantees unprecedented transparency, on the other hand, it makes it accessible by a vast number of subjects, which makes it difficult to exercise effective control over one's own data, implying tensions with the very principle of privacy by design and accountability, and with the principle of minimisation.

The immutability of the blockchain conflicts with the principle of storage limitation (Art. 5(1)(e) GDPR), the right to erasure (Art. 17 GDPR) and the right to rectify personal data (Art. 16 GDPR).

Another difficult point to resolve concerns the identification of the data controller and data processor[26] on the blockchain, given its decentralized and pseudo-anonymized character of the participants.

In truth, in the case of private and/or permissioned blockchains, the solution should be rather simple: in the first case, the data controller

[24] Cf. A.M. Gambino, C. Bomprezzi, op. cit., pp. 628 ff.

[25] For a more detailed discussion, Ivi, pp. 632 ff.

[26] The Data Controller (Art. 24 GDPR) is "the natural or legal person, public authority, agencies or other body" whose task is to determine the purposes, means and organization of the processing, and to designate the person(s) responsible for the processing. The Processor (Art. 28 GDPR), on the other hand, is the person who actually carries out the processing on behalf of the owner.

should be the infrastructure manager; in the second, the entity or entities that establish the access selection criteria[27].

In the case of public and/or permissionless blockchains, on the other hand, due to the highly decentralized system, this task is rather difficult[28]. On this point, among the many theses, two opposing ones are given below, in order to highlight the difficulties in resolving the issue: according to the first one[29], all nodes would be holders per se, because these correspond to the users who enter the data that are processed for a given purpose; for the second one instead, holder would be no one, because the individual nodes would not be able to carry out such a control. Then there is an intermediate way[30], according to which the controller could be identified each time in the entity that chooses to collect the personal data that are processed for a given purpose; the question, however, is still open.

4. Liability and Compensation Profiles

Further issues to be investigated regarding the relationship between blockchain technology and privacy include, undoubtedly, those related to compensation profiles, arising from the use of blockchain in the event of security deficits and violations of tracking systems (but also of data protection).

In fact, a regulatory framework is needed to highlight the best potential of blockchain technology, unveiling and overcoming its pitfalls and drawing out its liability[31]. One of the major questions that arises with reference to private law is precisely the assessment of the degree to which the blockchain's operating structure affects the area of civil liability, when damage is caused to users because of platform malfunctions[32].

One of the issues to be resolved relates to the choice between, on the one hand, the creation and preparation of an ad hoc remedial system capable of adapting to the rapid changes and developments in technology and, on the

[27] A.M. Gambino, C. Bomprezzi, *op. cit.*, p. 636.

[28] On this point, M. Giuliano, *La blockchain e gli smart contracts nell'innovazione del diritto nel terzo millennio*, cit., pp. 989 ff.: «*si pone, dunque, un problema di governance del sistema tecnologico, tenuto conto che la conformità al GDPR non riguarda la tecnologia, ma piuttosto come la tecnologia viene utilizzata per elaborare i dati personali*».

[29] Cautiously advanced, among others, by M. Giuliano, *op. cit.*, pp. 989 ff.

[30] Highlighted by A.M. Gambino, C. Bomprezzi, op. cit.

[31] M. Callegari, *Sostenibilità, supply chain e intelligenza artificiale*, cit., p. 1219.

[32] L. Buonanno, *La responsabilità civile nell'era delle nuove tecnologie: l'influenza della blockchain*, in «Resp. civ. prev.», V (2020), p. 1619. The A., in order to give an account of the incision that the topic of civil liability for damages caused by blockchain malfunctions represents, uses the expression "algorithmic civil liability".

other hand, the adaptation of the remedial and compensatory provisions already present within the legal system[33]. This point is still extremely controversial today. We will therefore make a proposal[34] in the knowledge that a clarifying legislative intervention is needed as soon as possible.

First, it is necessary to clarify one issue: the opinion that blockchain, even when subjected to stringent controls and checks, can be malfunctioning and characterized by design errors is almost universally shared. In support of this assertion, it suffices to recall that access (download) to the program to use the platform is frequently conditioned by the user's acceptance of and subscription to certain disclaimers by the platform operator[35] and by the user's choice of the remedy to be adopted in the event of system malfunctioning[36]. As has been pointed out by several parties, this aspect, linked to the automatisms inherent in this technology (including those, mentioned above, in the execution "of the agreement"), may give rise to a dangerous vacuum of protection for users, creating zones of irresponsibility and encouraging illicit uses of the system[37].

Consequently, in the event of damage resulting from errors in the design of the program or malfunctions of the blockchain technology, a hypothesis of strict liability could be configured for the managers of the platform. The managers, in fact, are perhaps the only subjects able to know the potentialities and limitations of the platform and, therefore, able to guarantee the maximum security of the technological service made available to users and to intervene in the event that this objective is not achieved[38]. The recourse to a system of strict liability is also justifiable by observing that, for such technologies, it is

[33] G. Bevivino, *Smart contract e "assestamenti" della disciplina generale del contratto,* cit., pp. 421 ff., who observes that «*Il limite di questa ricostruzione è, ad ogni modo, la tendenziale incompletezza del contratto che potrebbe dare vita alla mancanza assoluta di previsioni per determinate fattispecie patologiche che rimarrebbero, così, senza alcun tipo di protezione giuridica*».

[34] On this topic, M. Chierici, *Contratto di Blockchain as a Service: fondamenti teorici di una nuova prassi commerciale,* in «i Contratti», II (2022), pp. 197 ff.

[35] The operator's interest in protection, as observed by L. Buonanno, *La responsabilità civile nell'era delle nuove tecnologie,* cit., pp. 1619 ff., shows that the exploitation of platforms entails a number of potential risks for the user.

[36] M.F. Tommasini, *Lo smart contract e il diritto dei contratti,* in «Jus Civile», IV (2022), p. 855.

[37] F. Calisai, *Intelligenza artificiale e ambiente*, in «Giust. civ.», IV (2021), pp. 895 ff.

[38] M. F. Tommasini, *Lo smart contract e il diritto dei contratti,* cit., p. 855: «*poiché il funzionamento decentralizzato della blockchain implica che il gestore non abbia il controllo delle operazioni compiute (il che già esclude una responsabilità fondata sulla colpa), è ragionevole ritenere che un eventuale danno subito dagli utenti per il malfunzionamento della piattaforma gravi su chi poteva evitarlo. In capo al gestore-creatore della piattaforma può, dunque, ipotizzarsi una responsabilità di natura oggettiva che lo obbliga a risarcire i danni patiti dagli utenti*».

On this point see, among others, L. Buonanno, *La responsabilità civile nell'era delle nuove tecnologie*, cit., pp. 1619 ff.; F. Calisai, *Intelligenza artificiale e ambiente*, cit., pp. 895 ff.

often technically impossible or extremely difficult to prevent, identify or remove the malfunctioning of the system or machine[39].

In support of this, it should be noted that the European Parliament Resolution of 16 February 2017 on recommendations concerning civil law rules on robotics – but the discussion can be extended to technology based on distributed registries – already called on the Commission to consider introducing compulsory insurance mechanisms also for technologies based on distributed registries. The platform operator, as of today, seems to be the only entity that, through the instrument of insurance, would be able to incorporate risk within its cost-benefit analysis. One thinks of the planned ex ante payment of an insurance premium[40]: a system of compulsory insurance policies or ad hoc asset funds would make it possible to amortize, at least in part, the manager's liability[41].

Such a system, among other things, would also be able to bring about a collective benefit in terms of increasing users' confidence in the technology, incentivizing its use and, consequently, its improvement[42].

As a result of the above reflections, what emerges is that, although it is not clear what the future of regulation on the subject will be, there is a need to keep up-to-date and study the protection strategies to be adopted in the event of malfunctions of blockchain platforms, since there is a prospect of an intensification of the damage resulting from the use of such technologies and the violation of processed data. All this is, moreover, related to the need to achieve the widest possible involvement of the actors involved in this process, namely, on the one hand, the public and private regulators, but above all, on the other hand, the individuals themselves, who, by making use of such technologies, can grasp better than anyone else the future and possible cyber threats they entail.

[39] F. Calisai, *op. cit.*

Cf. F. Longobucco, *Il problema della soggettività giuridica delle macchine,* in *Diritto privato digitale,* a cura di E. Battelli, cit., pp. 33 ff.

[40] L. Buonanno, *La responsabilità civile nell'era delle nuove tecnologie,* cit. pp. 1621 ff.

[41] F. Calisai, *Intelligenza artificiale e ambiente,* cit., pp. 895 ff.

[42] In this regard, M.F. Tommasini, *Lo smart contract e il diritto dei contratti,* cit., pp. 855 ff, observed that, despite the considerable costs it would entail, compulsory insurance would have positive implications both for the operator, who can transform the risk of the service offered into a foreseeable cost of its activity, and for users, who would not incur the possible failure to realize their interests underlying the indemnity claim. This would determine, as a further reflex effect, a collective benefit in terms of user confidence in this type of market and its proper functioning.

BIBLIOGRAPHY

Alpa G., *Code is law: il bilanciamento dei valori e il ruolo del diritto*, in «Contratto e impresa», II (2021), p. 386 ff.

Annunziata F., *Blockchain and financial law: Fin Tech and crypto-assets*, in *Blockchain and public law. Global Challenges in the Era of Decentralisation*, a cura di O. Pollicino e G. De Gregorio, Elgar, Cheltenham 2021.

Attanasio C., *Inadempimento dello smart contract, sistema rimediale e tutela effettiva*, in «Rivista di diritto civile», IV (2024), pp. 719-743.

Battelli E. (con G. D'Ippolito e T. Pellegrini), *Diritto privato digitale*, Giappichelli, Torino 2024.

Bevivino G., *Smart contract e "assestamenti" della disciplina generale del contratto*, in «Contratto e impresa», II (2024), pp. 376-423.

Buonanno L., *La responsabilità civile nell'era delle nuove tecnologie: l'influenza della blockchain*, in «Responsabilità civile e previdenza», V (2020), pp. 1618-1631.

Billi A., *Blockchain e smart contract: commento all'attuale normativa italiana DLT e rilievi comparati*, in «Diritto e processo», XX (2020) p. 416 ff.

Caforio G., *La blockchain determina difficoltà nel determinare le regole da applicare e la giurisdizione competente*, in «Il Diritto industriale», II (2023), pp. 149-156.

Calisai F., *Intelligenza artificiale e ambiente*, in «Giustizia civile», IV (2021), pp. 895-910.

Callegari M., *Sostenibilità, supply chain e intelligenza artificiale*, in «Giurisprudenza italiana», V (2024), pp. 1211-1219.

Casarosa F., Comandé G., *Aspettando la NIS2: ovvero il diritto privato della cybersecurity*, in «Diritto dell'informazione», I (2024), pp. 29-53.

Cerrato S. A., Culasso F., Crocco E., *"Handle with care". Per una governance dell'intelligenza artificiale nell'impresa: rischi, tecniche di gestione, assetti cibernetici*, in «Rivista delle società», II-III 2023, pp. 371-414.

Chierici M., *Contratto di Blockchain as a Service: fondamenti teorici di una nuova prassi commerciale*, in «I Contratti», II (2022), pp. 197-219.

Cuffaro V., D'orazio R., Ricciuto V., *I dati personali nel diritto europeo*, Giappichelli, Torino 2019.

Di Gregorio V., *Intelligenza artificiale e responsabilità civile: quale paradigma per le nuove tecnologie?*, in «Danno e responsabilità», I (2022), pp. 51-63.

Gabrielli E., *La nozione di contratto e la sua funzione. Appunti sulla prospettiva di una nuova definizione di contratto*, in «Giustizia civile», I (2019), p. 299 ff.

Gambino A. M., Bomprezzi C., *Blockchain e protezione dei dati personali*, in «Diritto dell'informazione», III (2019), pp. 619-646.

Gambino A. M., Bomprezzi C., *Blockchain e criptovalute*, in *Fintech: diritti, concorrenza, regole – Le operazioni di finanziamento tecnologico*, a cura di G. Finocchiaro e V. Falce, Zanichelli, Torino 2019.

Giuliano M., *La blockchain e gli smart contracts nell'innovazione del diritto nel terzo millennio*, in «Diritto dell'informazione», VI (2018), pp. 989-1042.

Macrinici D., Cartofeanu C., Gao S., *Smart contract applications within blockchain technology: A systematic mapping study*, in «Telematics and Informatics», 2018, pp. 2337-2354.

Maugeri M., *Smart contract e disciplina dei contratti*, Il Mulino, Bologna 2021.

Poncibò C., *Il diritto comparato e la blockchain*, ESI, Napoli 2020.

Proietti G., *Definire l'indefinibile? I sistemi di intelligenza artificiale alla ricerca di un inquadramento sistematico*, in «Contratto e impresa», III (2024), pp. 882-925.

Proto M., *Questioni in tema di intelligenza artificiale e disciplina del contratto*, in *Il diritto nell'era digitale. Persona, mercato, amministrazione, giustizia*, a cura di R. Giordano, A. Panzarola, A. Police, S. Preziosi E M. Proto, Giuffrè, Milano 2022.

Rubechini P., *Criptoattività e blockchain*, in «Giornale di diritto amministrativo», II (2023), pp. 263-275.

Szabo N., *Formalizing and Securing Relationships on Public Networks*, in «First Monday», IX (1997).

Sirgiovanni B., *Lo "smart contract" e la tutela del consumatore: la traduzione del linguaggio naturale in linguaggio informatico attraverso il legal design*, in «Nuove leggi civili commentate», I (2023), pp. 214-226.

Tommasini M. F., *Lo smart contract e il diritto dei contratti*, in «Jus Civile», IV (2022), p. 831 ff.

Tosi E., *Responsabilità civile per illecito trattamento dei dati personali e danno non patrimoniale*, Giuffrè, Milano 2019.

Yeung K., *Regulation by Blockchain: The Emerging Battle for Supremacy between the Code of Law and Code as Law*, in «The Modern Law Review», II (2019), p. 207 ff.

GIUSEPPE MARIA MARSICO*

Artificial Intelligence: between market, regulation, subjectivity and risk allocation of responsibility§

Abstract

The undoubted difficulty of fully understanding, on a technical, dogmatic and practical-applicative level, the phenomenon of the use of Artificial Intelligence devices is at the basis of the multiplicity of reconstructions that have been proposed by those who have addressed the problem from the perspective of market and of the multifunctional system of civil liability. According to one reconstruction, an effective remedy to the "risk of autonomy" of AI would be to consider autonomous software agents as digital salespeople, calling the principal to account in the event of illegal harmful conduct. This thesis attributes a form of subjectivity to the software agent, exposing itself to criticism in terms of compatibility with the system of civil liability and subjectivity. The possibility of resorting to the attribution criterion provided for by art. 2052 of the Civil Code, thus equating the digital agent with the animal, leads to somewhat problematic outcomes: on the basis of this ontological diversity it would not be possible to operate either through analogical or extensive interpretation. Ai Act represents, from this perspective, a fundamental regulatory tool aimed at allocating risk efficiently and protecting the market.

Il contributo si propone di analizzare il ruolo del minore nell'ambiente digitale, e in particolare su Internet, luogo atterritoriale per eccellenza, dove chiunque ha la possibilità di interfacciarsi con un numero indefinito di utenti e contenuti. Si tratta senza dubbio di una straordinaria

* Pd.D. (c) Sapienza, University of Rome

§ This contribution develops the presentation at the conference 'Cybersecurity e prodotti: sfide e responsabilità nelle società interconnesse',29 November 2024, Italian Supreme Court.

opportunità per favorire lo sviluppo dell'identità di una persona, ma allo stesso tempo può presentare pericoli spesso sottovalutati, soprattutto quando i soggetti coinvolti sono persone particolarmente vulnerabili come i minori. L'articolo offre un'analisi sul trattamento dei dati dei minori, che è una delle principali innovazioni contenute nel Regolamento UE n. 2016/679 (GDPR) e alla quale il legislatore europeo ha dedicato nell'articolo 8 una specifica normativa sull'età del consenso. L'attenzione è rivolta alla questione dei sistemi di verifica dell'età nella dimensione online, poiché la verifica dell'età deve avvenire nel rispetto del suddetto regolamento e di alcuni principi, tra cui quello della minimizzazione dei dati. Pertanto, vengono illustrati alcuni esempi di metodi di verifica dell'età adottati dalle più importanti piattaforme online al fine di dimostrare la necessità di un equilibrio tra la sicurezza del minore e la tutela dei suoi diritti e libertà fondamentali.

1. Introductory remarks

There is no doubt that, over the past few years, doctrinal speculation on extra-contractual civil liability for damages caused by Artificial Intelligence systems has taken on dimensions of considerable importance. Notwithstanding the relative novelty of the appearance of the subject in the sphere of legal studies, it seems that the time is ripe, therefore, to start focusing attention on specific and delimited profiles of the subject under consideration. Very close and ontologically complex is the relationship between Artificial Intelligence systems, the market, algorithms and data. They appear directly and functionally affected by the relentless development of technological sciences, since AI systems perform multiple tasks based on predetermined instructions and the analysis of huge masses of information by means of algorithms.

Generally speaking, in terms of definition, the term 'algorithm' can be used to refer to the finite sequence of elementary operations, easily executable by a computer that, starting from a set of data (so-called input), produces another set of data (so-called output) that satisfy a pre-assigned set of requirements[1]». We refer, therefore, to the set of instructions, now increasingly complex, that guide the machine in the autonomous achievement of a given result.

[1] To mention only some of the most significant contributions in this field, and without any claim to completeness,, cfr. G. Alpa, *Diritto e Intelligenza Artificiale,* 2020, pp. 322 ff.; S. Infantino, *La responsabilità per danni algoritmici: prospettive europeo-continentali,* in Resp. Civ e prev., V (2019), pp. 1762 ff.; M. Costanza, *L'intelligenza artificiale e gli stilemi della responsabilità civile*, in Giur. it., (2019), pp. 1686 ff.; U. Ruffolo, *Intelligenza Artificiale, machine learning e responsabilità da algoritmo, ibidem,*

A very common example is the profiling activity, as is the use in the health sector, for the performance of surgical operations or the formulation of diagnoses. Similarly, in the judiciary, auxiliary systems are increasingly used to perform different activities, including research as well as decision-making.

Furthermore, the large-scale use of increasingly sophisticated and complex algorithms has made intelligent automated systems more efficient, as they are no longer based solely on initial programming code, but instead on machine learning methods.

Indeed, machine learning allows the algorithm to learn, improve and perfect itself throughout its 'life cycle', enabling the machine to increase its performance. The consequence of this is an increase in the use of data, also in the form of big data, from which comes the achievement of ambitious results very quickly, while at the same time making the human role increasingly marginal. Hence the consequence and also the risk that the role of the human person will become increasingly irrelevant and man will end up no longer being the beneficiary subject of technologies but, instead, a mere succumbing and passive object.

Indeed, much time has passed since man's life was dependent only on his abilities, physical, technical and intellectual.

If initially he could only rely on his own strength, or on animal strength, then came the first technologies and the refinement of new techniques, tools and materials, until today.

The development of computer science, algorithms and Artificial Intelligence systems as a whole are a clear and evident example of this, of which even the legal world, with reference to both public and private relations[2], is taking note in the twofold aspect of the need to mediate between the instances of technological and market development and, at the same time, provide adequate protection for individuals and their fundamental rights, in the face of phenomena long unknown to both jurisprudence and the legislature, weighing first of all the risks against the undoubted advantages.

pp. 1689 ff.; G. Finocchiaro, *Intelligenza artificiale e responsabilità*, in Contratto impr., II (2020), pp. 713 ff.; A. Fusaro, *Quale modello di responsabilità per la robotica avanzata? Riflessioni a margine del percorso europeo*, in *Nuova giur. civ. comm.*, VI (2020), pp. 1344 ff.

[2] Cfr., G. Sartor, F. Lagioia, *Le decisioni algoritmiche tra etica e diritto*, a cura di U. Ruffolo, Milano, 2020, pp. 84 ff.; G. Avanzini, *Decisioni amministrative e algoritmi informatici*, Editoriale Scientifica, Napoli, 2019; R. Cavallo Perin, *Ragionando come se la digitalizzazione fosse data,* in Diritto amministrativo, II (2020), pp. 305 ff.; E. Carloni, *I principi della legalità algoritmica. Le decisioni automatizzate di fronte al giudice amministrativo,* in Diritto amministrativo, II (2020), p. 273 ff.; G. Carullo, *Decisione amministrativa e intelligenza artificiale,* in Diritto dell'informazione e dell'informatica, III (2021), p. 431 ff.; S. Civitarese Matteucci, *Umano troppo umano. Decisioni amministrative automatizzate e principio di legalità,* in Diritto pubblico, I (2019), pp. 5 ff.; F. Costantino, *Rischi e opportunità del ricorso delle amministrazioni alle predizioni dei big data,* in Diritto pubblico, I (2019), pp. 43 ff.

Fundamental was the attainment of an adequate and mature level of technological development to give substance to the ideas prefigured by scientists, thus initiating the so-called 'Fourth Industrial Revolution'[3], as a phenomenon now characterised by a marked 'algocratic' matrix.

The ever-increasing 'autonomy' of machines poses dogmatic questions on the subject of responsibility and subjectivity, the solution to which can no longer be postponed, as it once was, and which come to the attention of scholars, both moral and ethical questions and legal questions, in the strict sense of the word, connected ex multis to the phenomenon of responsibility".

The ever-increasing 'autonomy' of machines poses dogmatic questions on the subject of responsibility and subjectivity, the solution to which is no longer, as it once was, procrastinable, as moral and ethical as well as legal questions, in the strict sense of the word, linked ex multis to the phenomenon of responsibility, come to the attention of the scholar.

First, there is the need to decide where and how to direct technological development, whether to adopt a techno-centric or human-centric approach[4], deciding, therefore, whether man, as the exclusive holder of constitutionally guaranteed fundamental rights, is to be the model to which technologies are to be benchmarked and to which they are to be made subservient, or, conversely, whether technology is to be at the centre and, as such, capable of challenging even some of the cornerstones of the protection of the human legal sphere, to the point of overcoming it.

At the centre of such a decision, however, must always be the cardinal legal value of our legal system of the dignity of the person and the constitutionally guaranteed rights to personal identity and non-discrimination, but also to transparency.

If one inclines towards the anthropocentric majority reconstruction, the question then follows as to which activities it is possible or appropriate to delegate to an automated system.

Further problems of a legal nature are in fact connected with this, such as the possibility of attributing legal subjectivity to it and, therefore, the ownership of subjective legal situations, with the associated liability profiles.

[3] Cfr., L. D'Avack, *La rivoluzione tecnologica e la nuova era digitale: problemi etici,* in U. Ruffolo (ed), Intelligenza Artificiale. Il diritto, i diritti, l'etica, Milano, 2020.

[4] A. Simoncini, *L' algoritmo incostituzionale: intelligenza artificiale e il futuro delle libertà,* in *BioLaw Journal*, I (2019), pp. 63 ff.

2. Allocation of liability risk and the market: do we really need to rethink dogmatic categories of subjectivity and liability?

Borrowing a term that has come into widespread use, it would indeed be difficult to deny that the subject of civil liability for damages arising from the use of Artificial Intelligence[5] devices has been a trend topics in civil law literature for several years now, so copious in number and dense in content are the writings that have addressed the issue, often in the context of discourses that are also attentive to its possible interference with the prospects of future development of the institution of tort law[6].

The reconstructive hypotheses and argumentative proposals that have been elaborated in this way seem to have explored every possible area of the debate, also dwelling on practical application problems in the sector, such as those concerning the use of Artificial Intelligence applications in medicine and the impact of the same on the liability of the healthcare professional who avails himself of them and, of course, the subject of liability in respect of the use of driverless cars and other remotely piloted vehicles: so that, and all the more so in a context that is still lacking in significant contributions from case law, a further intervention on the subject would seem destined to resolve itself into a review, at this point of a merely reconnaissance nature, of the various opinions that have stratified on the subject.

Therefore, it is necessary, at this stage of development of a discourse that is naturally still to be completed – and not only in terms of regulatory evolution – to verify whether the strategy envisaged by the Proposal for a Directive on the subject of AI Liability is capable of suggesting to the interpreter certain considerations even in the current state of the law and, therefore, based exclusively on the interpretation of the current and, indeed, already complex regulatory framework[7].

What has just been hypothesised is, on the other hand, the way of dealing with the subject that allows one to grasp in the clearest terms the grafts on the evolution of civil liability as a whole: not only as regards the level of the

[5] V. Di Gregorio, *Intelligenza artificiale e responsabilità civile: quale paradigma per le nuove tecnologie?*, in Danno e responsabilità, I (2022), pp. 51 ff.

[6] G. Sartor, *L'intenzionalità dei sistemi informatici e il diritto, in Rivista trimestrale di diritto e procedura civile*, I (2003), pp. 23 ff.; Id., *Gli agenti software: nuovi soggetti del ciberdiritto?,* in Contr. impr., II (2002), pp. 465 ff.; G. Taddei Elmi, *I diritti dell'intelligenza artificiale tra soggettività e valore: fantadiritto o ius condendum?,* in L. Lombardi Valluri (ed), *Il Meritevole di tutela,* Giuffrè, Milano, 1990, pp. 685 ff.

[7] A. Mantelero, *Artificial Intelligence, dati e diritti: spunti di riflessione per i regolatori*, in P. Bertoli, F. Ferrari, G. Ripamonti, G. Tiberi (ed), *Data protection tra Unione europea*, Italia e Svizzera, 2019, pp. 8 ff.

criteria for imputation of the indemnity obligation, but also as regards the role that the institute of tort law is able to play within a phase of scientific and technological development which, to a first approximation, would appear to be truly incapable of being read according to the canons accredited up to now and disciplined by the regulatory techniques currently made available by the regulatory system. To the point that, as has been observed quite recently, Artificial Intelligence should not be seen so much as 'the object of advanced and innovative research, as the context within which legal research must rewrite its categories'; in other words, there would not be 'a legal problem of Artificial Intelligence', since the various conflicts of interest are instead located 'to a very great extent, not to say in all cases, in the society of the new technology[8]».

Analysing the close relationship between AI and the protean institution of transparency is useful both to clarify which characteristics an automated intelligent system should respect in order to allow human control, and to identify the possible subjects responsible for any damage. It has, in fact, been pointed out in doctrine and case law that these systems should be as far as possible controllable by humans, with a view to affirming their exclusive subjection to a merely instrumental use.

On the merits, however, one must ask oneself whether it is correct and opportune to have recourse to the common civil liability regimes or to refer to a different legislation; and also to consider whether to invest AI with legal subjectivity and thus the possibility of an automated system being directly liable for damages caused[9].

On this last point, although there are different positions, in truth, the legislator has already definitively made a choice, abandoning (at least for now) the hypothesis of so-called electronic subjectivity. On the other hand, the existing positions in doctrine are generally negative, albeit with a few favourable exceptions, so that Artificial Intelligence is not recognised, at least to date, as having the typical requisites of a legal subject but rather those of an object to which, for example, the typical protections of property, copyright and patent institutes can be attributed.

The decision to abandon the idea of the 'electronic' personality of AI systems has, by now, also matured in the intentions of the European legislator who, with a changed perspective compared to the recent past, has affirmed that 'human beings must always be responsible'.

[8] G. Taddei Elmi, *I diritti dell'intelligenza artificiale tra soggettività e valore, fantadiritto o ius condendum?*, cit., pp. 705-707.

[9] G. Taddei Elmi, S. Marchiafava, *Sviluppi recenti in tema di Intelligenza Artificiale e diritto: una rassegna di legislazione, giurisprudenza e dottrina,* in Riv. it. inf. e dir., II (2022).

Accompanied by the extensive and now mature intervention of the European legislator[10], there arises the age-old querelle of the applicability, at least in itinere, of a Euro-Union liability law. In this hermeneutic perspective, it seems relevant not only to examine the compatibility between the different disciplines and AI systems, but also to assess whether or not it is appropriate to refer to a common liability regime in a highly fragmented domestic framework.

With regard to the complex analysis of the problematic profiles of the individual liability regimes with respect to damages derived from the use of Artificial Intelligence, it is possible to affirm the exclusion of the applicability of Article 2043 of the Civil Code[11], which presupposes, at least with reference to domestic law, the action of a subject, which, as argued above, cannot be understood to be an Artificial Intelligence system.

On the other hand, a different reconstruction is required for the regimes of indirect or strict liability under Articles 2047 et seq. of the Civil Code, according to which a person who is in a particular relationship of proximity with the person or thing from which the damage originated is held liable. Indeed, the possibility of tracing the liability system relating to IA to one of the many figures provided for by the Civil Code is affected by the diversity of the IA systems themselves to which, depending on the case, different forms of liability could be traced. Among other things, this would require an excessive effort on the part of the judge, such as to potentially mislead him, with significant consequences also in terms of economic and technological development, starting with the distrust of producers and users.

In doctrine, in particular, the possibility of referring to the discipline on the subject of damage caused by the circulation of vehicles, provided for by Article 2054 of the Civil Code, has also been considered; also due to the fact that the case of self-driving cars has been one of the primary hypotheses in which

[10] U. Pagallo, *The law of robots. Crimes, contracts and torts*, New York, 2013, pp. 128 ff.; G. Teubner, *Soggetti giuridici digitali? Sullo status privatistico degli agenti software autonomi,* ESI, Napoli, 2019, a cura di P. Femia; as well as the provisions of art. 2050 c.c. concerning damage caused in the course of carrying out dangerous activities (in this sense L. Coppini, *Robotica e intelligenza artificiale: questioni di responsabilità civile,* in Pol. Dir., IV (2018), p. 735, il quale evidenzia i «nuovi pericoli di fallibilità del prodotto »; M. Scialdone, *Il diritto dei robot: la regolamentazione giuridica dei comportamenti non umani*, in E. Pietrafesa, F. Marzano, T. Medici (a cura di), *La rete e il fattore C: Cultura, Complessità, Collaborazione*, vol. II, Roma, Stati Generali dell'Innovazione, 2016, pp. 76 ff.; G. D'Alfonso, *Il regime di responsabilità da cose in custodia tra questioni tradizionali e "responsabilità da algoritmo",* in *EJPLT*, I (2022), pp. 82 ff.; A. bertolini, *Robots as Products: The Case for a Realistic Analysis of Robotic Applications and Liability Rules,* in Law, Innovation and Technology, 5, II (2013), pp. 214 ff.

[11] G. Calabresi, E. Al Mureden, *Driverless cars: intelligenza artificiale e futuro della mobilità*, Il Mulino, Bologna, 2021.

the question of civil liability for AI damage has arisen. See, in this regard, how one of the solutions put forward to make up for the current lack of a perfectly applicable regulation within the legal system comes from this very sector. The thesis is based on the provision of an obligation to take out an insurance policy for persons who make use of Artificial Intelligence systems in various capacities, comparable to the existing obligation for car owners[12].

In order to ensure adequate protection for those harmed by an AI system, the possibility of applying – extensively – the Product Liability Directive has also been raised[13], on the basis of the assumption that AI, like the machines in which it is integrated, is classifiable in the category of 'products' even when it lacks materiality.

It seems evident that the provision of an objective system for attributing liability would make it possible to guarantee effective protection for the injured party, overcoming the critical issues arising from the need to identify a person to whom to attribute behaviour that is at least culpable. However, at present the Product Liability Directive, as also transposed into national law, is applicable only to damage generated by AI 'products', thus requiring an interpretive adjustment due to the difficulty of including all systems and their future development, an objective taken into consideration by the European Commission.

Thus, the current regulatory framework is, as of today, blurred, uncertain and fragmented, giving rise to not a few problems relating to the protection of individuals, at the application level, given the uncertain and debated possibility of referring to national non-contractual liability rules, on the one hand, and to product liability[14] rules, on the other.

[12] A. Davola, R. Pardolesi, *In viaggio con il robot: verso nuovi orizzonti della r.c. auto ("driveless")?*, in Danno e resp., V (2017), pp. 625 ff.; A. Bertolini, E. Palmerini, *Regulating robotics: A challenge for Europe, Publications Office of the EU Parliament, Bruxelles*, 2014, pp. 185-186. The use of an insurance system involving those involved in the whole supply chain, from development to the use of AI, had already been considered by a study commissioned by the European Parliament, entitled "A common EU approach to liability rules and insurance for connected and autonomous vehicles", February 2018, available at the following link: EPRS_STU(2018)615635_EN.pdf (europa.eu). Report problems arising from the use of insurance, such as identifying parameters for setting insurance premiums, L. Coppini, *Robotica e intelligenza artificiale: questioni di responsabilità civile,* cit., p. 737.

[13] In tal senso si veda N.F. Frattari, *Robotica e responsabilità da algoritmo. Il processo di produzione dell'intelligenza artificiale,* in Contr. impr., I (2020), pp. 458 ff.

[14] P.G. Chiara, *Sistemi intelligenti autonomi e responsabilità civile: stato dell'arte e prospettive nell'esperienza comunitaria, i*n Diritto ed economia dell'impresa, I (2020), p. 128.

3. Brief concluding remarks: between the risk-based approach of the AI Act and balancing of interests

In order to mitigate the multifaceted application problems arising from high-risk AI systems placed on the market or put into service, and to ensure a high level of reliability, it seemed appropriate to apply certain mandatory and mandatory requirements to high-risk AI systems, taking into account the intended purpose and context of use of the AI system and in accordance with the risk management system to be established by the supplier.

In this sense, an approach based on the economic analysis of the law seems useful[15], functional to an optimal allocation of the negative externalities arising from the occurrence of liability (so-called Paretian optimum), in order, moreover, from a legal policy perspective, not to stifle technological development. Measures taken by providers to comply with the mandatory requirements of this regulation should take into account the generally acknowledged state of the art in AI and be proportionate and effective to achieve the objectives of this regulation[16]. On the basis of the new legislative framework, as clarified in the Commission Communication 'The "Blue Guide" to the implementation of EU product legislation 2022', Union harmonisation legislation can normally only be applicable to a product, as it can only be made available or put into service if it complies with all existing Union harmonisation legislation.

The hazards of IA systems covered by the requirements of the Regulation concern different aspects than existing Union harmonisation acts and therefore the requirements of this Regulation would complement the existing body of Union harmonisation acts.

The regulation aims to ensure that Artificial Intelligence systems placed on the European market are safe and respect the fundamental rights and values of the European Union. The AI Act also aims to stimulate investment and innovation in the Artificial Intelligence sector and facilitate the development of a single market for lawful, secure and trustworthy AI applications, positioning the EU as a major global player. Different national rules in each EU Member State could, in fact, lead to fragmentation of the internal market, threatening the EU's competitive advantage, and decrease legal certainty for operators developing, importing or using AI systems within the EU. The text of

[15] U. Cooter, U. Mattei, R. Pardolesi, T. Ulen, *Il Mercato delle Regole*, Bologna, Il Mulino, 1999, pp. 294 ff.

[16] F. Faini, *Il diritto nella tecnica: tecnologie emergenti e nuove forme di regolazione,* in *federalismi.it,* XVI (2020), p. 112. On the topic of "comprehensibility" of decisions based on algorithmic technologies that interfere with the personal sphere, v. A. Simoncini, *Il linguaggio dell' intelligenza artificiale e la tutela costituzionale dei diritti*, in Rivista AIC, II (2023), pp. 34 ff.

the regulation approved by the EU Parliament confirms the risk-based approach already adopted in the previous versions circulated in recent months. Thus, different obligations are imposed according to the level of risk that each of the categories into which AI systems can be classified may pose to the rights and freedoms of individuals. The approach followed by the Regulation is, therefore, that the greater the risk, the stricter the rules to be applied (the so-called risk-based approach). The risk-based approach is embodied in a threefold categorisation of Artificial Intelligence: only systems that fall into the unacceptable risk category are excluded; High-risk systems have minimum and rather general requirements, whereas low-risk systems are not regulated.

The classification of AI systems is based on identified risks to the rights and freedoms of individuals. In particular, AI systems are classified into the following categories defined according to their potential risks: (a) unacceptable risk; (b) high risk; (c) low or minimal risk. Of particular importance in this perspective are certain uses of AI systems that pose a serious threat to fundamental rights.

Therefore, the risk associated with such systems is considered unacceptable by the regulation and their use will be prohibited within the EU. The AI Act mentions among such systems, for instance, those that can be used for the purpose of cognitive-behavioural manipulation of the user (e.g. toys that use voice assistants to encourage dangerous behaviour in minors); so-called 'social scoring' or 'social scoring' systems; and 'social behavioural' systems that can be used for the purpose of encouraging dangerous behaviour in minors. social scoring' systems, i.e. those that score individuals on the basis of their behaviour, thus influencing their access to services, employment, or other opportunities; those that recognise emotions, when used in the workplace and in education; real-time biometric identification systems in publicly accessible spaces; the biometric categorization systems of natural persons used to extract and process sensitive data (such as sexual orientation or religious beliefs) and the systems aimed at the so-called "predictive policing" and used to make predictions of future crimes. Secondly, the AI Act aims to introduce a series of provisions that aim to ensure that high-risk AI systems are subject to rigorous requirements and obligations before being introduced or used in the EU. These systems include a wide range of devices, such as those used in critical infrastructure, medical devices, as well as biometric identification, categorization and emotion recognition systems. In light of the potential risk[17] associated with such applications, the obli-

[17] A. Simoncini, *L' algoritmo incostituzionale: intelligenza artificiale e il futuro delle libertà*, in *BioLaw Journal*, 2019, pp. 63 ff.; the need to balance risks and positive effects on the exercise and protection of rights and freedoms, v. F. DONATI, Diritti fondamentali e algoritmi nella proposta di regolamento sull' Intelligenza Artificiale, in A. Pajno, F. Donati, A. Perrucci (eds.), *Intelligenza artificiale e diritto: una rivoluzione?*, vol. 1, cit., pp. 112 ff.

gations imposed include the adoption of risk management systems, compliance with requirements relating to the quality of the datasets used (which must be high), as well as the adoption of technical documentation and record keeping, transparency and provision of information to users, as well as prescribing the adoption of adequate levels of accuracy, robustness and cybersecurity. The AI Act also requires that operators of high-risk AI systems carry out a fundamental rights impact assessment and that this assessment be carried out before such systems are placed on the market.

The recipients of obligations relating to high-risk systems are not only AI system providers, but also other actors along the AI value chain (e.g. importers, distributors, authorized representatives), on whom proportionate obligations are imposed in order not to undermine technological development and innovation.

The AI Regulation also aims to provide specific obligations for low-risk AI systems. This category includes most AI systems. The Regulation clarifies that the production and use of Artificial Intelligence systems that present only a limited risk to the rights and freedoms of individuals will be subject to simple transparency obligations. In particular, article 52 of the AI Act establishes, in relation to systems that interact with individuals, emotion recognition and biometric categorization systems not included among those prohibited, as well as systems that generate or manipulate so-called content "deep fake". In this context, it is envisaged the obligation to inform the user that he is interacting with an Artificial Intelligence system or of the fact that a particular content has been created through intelligence, in order to allow the user to use technology in an informed and conscious way.

In light of the genesis of the Regulation, which has its roots in the Guidelines for Trustworthy AI and in the White Paper on AI, it can reasonably be stated that the legal interests underlying the distinction between AI systems with unacceptable risk, risk and low risk are the fundamental rights of the individual. The gradation of risk, which guides the identification of the three regulatory categories just mentioned, is conducted by the Euro-unitary legislator on the basis of an ex ante assessment of the danger of the different AI systems with respect to the protection of fundamental rights.

In fact, the discipline identifies only typical cases and refrains from providing an overly detailed taxonomy of AI applications, which would soon be superseded by the rapid pace of innovation. The long-standing debate on the definition of high-risk systems, currently listed in Annex III of the AI Act, and on the opportunity to conduct ex post assessments of the risks that have emerged, reflect the difficulty of finding a compromise between legal certainty and general and future scope of a regulation that addresses a rapidly evolving te-

chnological context. This compromise is even more difficult to find in the face of the self-adapting and self-feeding characteristics of AI, which make any regulatory intervention shortsighted, even that resulting from the most targeted empirical exercise of regulatory sandboxes[18]. In order to ensure flexibility in the above categorization, the discipline also identifies a general category of AI systems, the so-called «general purpose» AI systems, which – as will be analysed in the following – are «intended by the supplier to perform generally applicable functions» and «may be used in a plurality of contexts and be integrated into a plurality of other AI systems». Compared to these, the AI regulation provides for the application of only some requirements for high-risk systems. These will be determined by a delegated act based, inter alia, on a specific impact assessment. The differentiated regimes of IA systems thus framed are intended to be applied transversely across the whole euro-unitary single market. This scope is consistent with the reticular physiognomy of digital markets where Artificial Intelligence systems are widely used. However, the declared horizontal approach of the legislation is likely to be undermined – in practice – both at national level and on the market, by what can be called a «limited» horizontal approach of the European regulation on Artificial Intelligence. Although the AI Act aims to provide a uniform horizontal regulatory framework for innovation in the European Artificial Intelligence market, it does give Member States substantial scope for choice in implementing many of the requirements laid down, which remain rather general and vague. Indeed, the implementation of requirements (especially for high-risk systems) is left to the processes of self-assessment and self-certification by enterprises – that is to say, AI system producers and users – , Fuelling potential divergences if not real inconsistencies in the European AI market[19].

In the absence of clear guidelines, the different policies for implementing regulatory requirements by individual market players could in turn undermine the cross-sectoral nature of the proposed legislative text. Furthermore, the regulation allows Member States, in recital 68, to waive new provisions for exceptional reasons relating to health, public safety or intellectual property.

The intention of the European legislator is that these potential atomistic drifts at primary regulatory level should be contained through soft law interventions by a new European authority responsible for overseeing AI: the

[18] A. Mantelero, *Artificial Intelligence, dati e diritti: spunti di riflessione per i regolatori,* in P. Bertoli, F. Ferrari, G. Ripamonti, G. Tiberi (a cura di), *Data protection tra Unione europea*, Italia e Svizzera, 2019, pp. 8 ff.

[19] On the topic of "comprehensibility" of decisions based on algorithmic technologies that interfere with the personal sphere, v. A. Simoncini, *Il linguaggio dell'intelligenza artificiale e la tutela costituzionale dei diritti,* in Rivista AIC, II (2023), pp. 34 ff.

European IA Committee. The advisory and regulatory tasks of the Committee and its collaboration with other supervisory authorities should ensure a uniform application of the rules in all Member States, While paying attention to the sectoral governance instances of the regulatory requirements introduced in the field of Artificial Intelligence, arising from the specificities of different market areas. While it is true, in light of the above, that for the purposes of the AI regulation all markets must be considered equal, in light of the proposed categorizations, it seems that some will be «more equal than others» in point of IA discipline. Indeed, although it is quite ambitious, the AI Act does not provide for holistic regulation and does not aim to regulate AI systems in every market sector. On the contrary, the horizontal approach of IA law is limited to its very objective, namely the protection of fundamental rights of European consumers, codified in relevant European regulatory documents such as the Nice Charter.

This limitation of the interests and objectives pursued by IA law means that those IA systems used in markets which have – at least at first sight – a less invasive impact on fundamental rights are excluded from the scope of its provisions of consumers/citizens[20].

From this brief reconstruction, it seems clear that the EU's approach to AI is based on trust, with the aim of enhancing research and industrial capacity while ensuring security and fundamental rights.

To promote innovation, in fact, as already mentioned, the AI Act provides for the creation of sandboxes, spaces for experimentation. These sandboxes allow companies, including small and medium-sized enterprises and start-ups, to experiment with AI solutions in realistic environments, while enjoying temporary regulatory exemptions.

At the same time, AI models specifically designed for research are also exempted from certain restrictions, to encourage – in a perspective of economic analysis of law – greater freedom in research and development. In addition, the AI Act provides for targeted exceptions to facilitate SMEs' adaptation to the new legislation, taking into account their limited resources compared with large players in the sector.

These regulatory sandboxes, in an interdisciplinary perspective as regards the law of the economy and financial markets, beyond the level of reporting obligations, Implement an additional front of collaboration between national supervisors and the European Committee for Artificial Intelligence. The objective of supervision is to ensure that controlled trials on similar technological applications are treated equally by different national competent au-

[20] F. Faini, *Il diritto nella tecnica: tecnologie emergenti e nuove forme di regolazione*, in *federalismi.it,* (2020), p. 112.

thorities. The testing of fintech applications, in particular, should be governed by the principle «same activity, same risk, same rules», which requires financial authorities, on the one hand, not to treat similar technologies differently and, on the other, «transplanting» the regulatory solutions developed within a specific sandbox in other application areas. The AI Act therefore encourages (c.d. inducement), from a Law & Economics perspective, cooperation between the competent national authorities, in full respect of the supervisory and corrective powers of the authorities supervising the sandbox; on the one hand, with the aim of stimulating investment and innovation in the field of Artificial Intelligence and facilitating the development of a single market for legal, safe and reliable AI applications, positioning the EU as a major global player; on the other hand, by carrying out a careful assessment of the impact on fundamental rights, or Fundamental Rights Impact Assessment (FRIA), such as privacy and environmental[21].

[21] G. Alpa, *Diritto e Intelligenza Artificiale*, 2020, pp. 322 ff.

BIBLIOGRAPHY

Alpa G., *Diritto e Intelligenza Artificiale*, Il Mulino, Bologna 2020, pp. 322 ff.

Avanzini G., *Decisioni amministrative e algoritmi informatici*, Editoriale Scientifica, Napoli 2019.

Bertolini A., *Robots as Products: The Case for a Realistic Analysis of Robotic Applications and Liability Rules*, in «Law, Innovation and Technology», V II (2013), pp. 214 ff.

Bertolini A., Palmerini E., *Regulating robotics: A challenge for Europe*, Publications Office of the European Parliament, Bruxelles 2014, pp. 185-186.

Calabresi G., Al Mureden E., *Driverless cars: intelligenza artificiale e futuro della mobilità*, Il Mulino, Bologna 2021.

Carloni E., *I principi della legalità algoritmica. Le decisioni automatizzate di fronte al giudice amministrativo*, in «Diritto amministrativo», II (2020), pp. 273 ff.

Carullo G., *Decisione amministrativa e intelligenza artificiale*, in «Diritto dell'informazione e dell'informatica», 3 (2021), pp. 431 ff.

Cavallo Perin R., *Ragionando come se la digitalizzazione fosse data*, in «Diritto amministrativo», II (2020), pp. 305 ff.

Chiara P. G., *Sistemi intelligenti autonomi e responsabilità civile: stato dell'arte e prospettive nell'esperienza comunitaria*, in «Diritto ed economia dell'impresa», I (2020), p. 128 ff.

Civitarese Matteucci S., *Umano troppo umano. Decisioni amministrative automatizzate e principio di legalità*, in «Diritto pubblico», I (2019), pp. 5 ff.

Coppini L., *Robotica e intelligenza artificiale: questioni di responsabilità civile*, in «Politica del diritto», IV (2018), pp. 735 ff.

Cooter U., Mattei U., Pardolesi R., Ulen T., *Il mercato delle regole*, Il Mulino, Bologna 1999, pp. 294 ff.

Costanza M., *L'intelligenza artificiale e gli stilemi della responsabilità civile*, in «Giurisprudenza italiana», (2019), pp. 1686 ff.

Costantino F., *Rischi e opportunità del ricorso delle amministrazioni alle predizioni dei big data*, in «Diritto pubblico», I (2019), pp. 43 ff.

D'alfonso G., *Il regime di responsabilità da cose in custodia tra questioni tradizionali e "responsabilità da algoritmo"*, in «EJPLT», I (2022), pp. 82 ff.

D'avack L., *La rivoluzione tecnologica e la nuova era digitale: problemi etici*, in *Intelligenza artificiale. Il diritto, i diritti, l'etica*, a cura di U. Ruffolo, Giuffrè, Milano 2020.

Davola A., Pardolesi R., *In viaggio con il robot: verso nuovi orizzonti della r.c. auto ("driverless")?*, in «Danno e responsabilità», V (2017), pp. 625 ff

Di Gregorio V., *Intelligenza artificiale e responsabilità civile: quale paradigma per le nuove tecnologie?*, in «Danno e responsabilità», I (2022), pp. 51 ff

Donati F., *Diritti fondamentali e algoritmi nella proposta di regolamento sull'intelligenza artificiale*, in *Intelligenza artificiale e diritto: una rivoluzione?*, vol. I, a cura di A. Pajno, F. Donati, A. Perrucci, Il Mulino, Bologna 2022, pp. 112 ff.

Faini F., *Il diritto nella tecnica: tecnologie emergenti e nuove forme di regolazione*, in «federalismi.it», XVI (2020), p. 112 ff.

Finocchiaro G., *Intelligenza artificiale e responsabilità*, in «Contratto e impresa», II (2020), pp. 713 ff.

Frattari N.F., *Robotica e responsabilità da algoritmo. Il processo di produzione dell'intelligenza artificiale*, in «Contratto e impresa», I (2020), pp. 458 ff.

Fusaro A., *Quale modello di responsabilità per la robotica avanzata? Riflessioni a margine del percorso europeo*, in «Nuova giurisprudenza civile commentata», (2020), pp. 1344 ff.

Infantino S., *La responsabilità per danni algoritmici: prospettive europeo-continentali*, in «Responsabilità civile e previdenza», V (2019), pp. 1762 ff.

Mantelero A., *Artificial Intelligence, dati e diritti: spunti di riflessione per i regolatori*, in *Data protection tra Unione europea, Italia e Svizzera*, a cura di P. Bertoli, F. Ferrari, G. Ripamonti, G. Tiberi, Giappichelli, 2019, pp. 8 ff.

Pagallo U., *The law of robots. Crimes, contracts and torts*, Springer, New York 2013, pp. 128 ff.

Ruffolo U., *Intelligenza artificiale, machine learning e responsabilità da algoritmo*, in «Giurisprudenza italiana», (2019), pp. 1689 ff.

Sartor G., *Gli agenti software: nuovi soggetti del ciberdiritto?*, in «Contratto e impresa», II (2002), pp. 465 ff.

Sartor G., *L'intenzionalità dei sistemi informatici e il diritto*, in «Rivista trimestrale di diritto e procedura civile», I (2003), pp. 23 ff.

Sartor G., Lagioia F., *Le decisioni algoritmiche tra etica e diritto*, in *Le decisioni algoritmiche tra etica e diritto*, a cura di U. Ruffolo, Giuffrè, Milano 2020, pp. 84 ff.

Scialdone M., *Il diritto dei robot: la regolamentazione giuridica dei comportamenti non umani*, in *La rete e il fattore C: Cultura, Complessità, Collaborazione*, vol. II, a cura di E. Pietrafesa, F. Marzano, T. Medici, Stati Generali dell'Innovazione, Roma 2016, pp. 76 ff.

Simoncini A., *L'algoritmo incostituzionale: intelligenza artificiale e futuro delle libertà*, in «BioLaw Journal», (2019), pp. 63 ff.

Simoncini A., *Il linguaggio dell'intelligenza artificiale e la tutela costituzionale dei diritti*, in «Rivista AIC», II (2023), pp. 34 ff.

Taddei Elmi G., *I diritti dell'intelligenza artificiale tra soggettività e valore: fantadiritto o ius condendum?*, in *Il meritevole di tutela*, a cura di L. Lombardi Vallauri, Giuffrè, Milano 1990, pp. 685 ff.

Taddei Elmi G., Marchiafava S., *Sviluppi recenti in tema di intelligenza artificiale e diritto: una rassegna di legislazione, giurisprudenza e dottrina*, in «Rivista italiana di informatica e diritto», II (2022).

Teubner G., *Soggetti giuridici digitali? Sullo status privatistico degli agenti software autonomi*, ESI, Napoli 2019, a cura di P. Femia.

RACHELE RANIERI*

Jurisprudence and Artificial Intelligence: the Romeo and Juliet of law§

Abstract

Love, passion, temptation, sorrow and death. Aspects, all, that express the particular and ambiguous relationship that exists, nowadays, between the world of law and Artificial Intelligence. The allure of new discovery and progress, in a society that is constantly evolving, has gone so far as to arouse the idea, previously unimaginable even in the deepest dreams, of a world in which the figure of «robot judge» is replaced by that of «human judge». But, a system of predictive justice, if it seems to «caress» the judicial apparatus with a safe and protective hand, at the same time, nevertheless, hides in the deep grave risks and dangers that could lead to a final loss of substantial Justice. The challenge of today and tomorrow is to create a system in which machine and man help each other and work together for a better order.

Amore, passione, tentazione, sofferenza e morte. Aspetti, tutti, che esprimono quel particolare e ambiguo rapporto che intercorre, ad oggi, tra il mondo del diritto e l'Intelligenza Artificiale. La seduzione della nuova scoperta e del progresso, in una società in continuo divenire, si è spinta a tal punto da suscitare l'idea, prima inimmaginabile neanche nei sogni più reconditi, di un mondo in cui la figura di «giudice robot» si sostituisca a quella di «giudice uomo». Ma un sistema cd di giustizia predittiva, se in apparenza sembrerebbe «accarezzare» l'apparato giurisdi-

* PhD student in "Diritto Romano, Teoria Degli Ordinamenti e Diritto Privato del Mercato" (Curriculum in Diritto Privato del Mercato) Sapienza, University of Rome.

§ This contribution develops the presentation at the conference 'Cybersecurity e prodotti: sfide e responsabilità nelle società interconnesse',29 November 2024, Italian Supreme Court.

zionale con una mano sicura e di protezione, allo stesso tempo, tuttavia, cela nel profondo gravosi rischi e pericoli che potrebbero sfociare in una definitiva perdita della Giustizia sostanziale. La sfida di oggi e di domani è quella di creare un sistema in cui la macchina e l'uomo traggano beneficio l'un l'altro e lavorino insieme per un ordinamento migliore.

1. The die is cast: a first approach to predictive justice

In a world and society that is constantly evolving technologically, the field of law, as well as that of process, are inevitably forced to come to terms with the revolutionary potential and risks associated with the use of Artificial Intelligence. As a preamble (which, in reality, might sound like an excessive mistrust of progress), what some philosophers had predicted may be considered to have come true, namely that «technology, understood as the universe of technologies and the rationality that presides over their functionality and efficiency, is no longer a means at man's disposal, but is the environment within which the Man too undergoes modification»[1]. In approaching the study of such a phenomena, it emerges how, in legal-as well as technical terms-providing a definition of the term «Artificial Intelligence» (AI)[2] is a complex operation. In recent times, it was the European legislator, with Regulation No. 1689/2024, the so-called Artificial Intelligence Act (AI Act or AIA), who was among the first to provide a normative definition.

To speak of Artificial Intelligence within the justice sector implies referring to an indeterminate, and perhaps as yet unimaginable, multiplicity of potentially usable machines, tools and software. In the process of modernizing society, AI has been playing a primary role for years, as it is evident with regard to the services' digitalization, the databases' creation, systems enabling the early identification of high-risk or high-priority cases, enabling judicial offices to allocate resources more efficiently, and the digitalization of appeals. Artificial Intelligence has, therefore, officially entered what Satta called the «mystery of the trial»[3].

[1] Argument put forward by A. Galimberti., *Psiche e techne. L'uomo nell'età della tecnica,* Feltrinelli, Milano 2016.

[2] The paternity of the term is unanimously attributed to the mathematician John McCarthy who in 1956 defined Artificial Intelligence as 'the science and engineering of building intelligent machines'. However, many see in A. Turing the father of intelligent computers and what would later become Artificial Intelligence.

[3] S. Satta., *Il mistero del processo,* in «Riv. dir. proc.», 1949, p. 273 ss. (text of the lecture given at the University of Catania on 4 April 1949, later republished in *Soliloqui e colloqui di un giurista,* Ilissio, Nuoro 1968, pp. 3-18 and in *Il mistero del processo*, Piccola Biblioteca Adelphi, Milano 1994, pp. 11-37).

Here we would like to draw attention to a specific application hypothesis of AI in the field of law, namely the transformation of the judicial phase into a mechanism of pure predictive justice. Such a system promotes the possibility of predicting procedural outcomes, i.e. the judicial decision, through a probabilistic calculation in which the scientific method is applied to the law. In the field of predictivity, indeed, AI is a valuable resource as it can analyze huge amounts of data quickly and efficiently in order to derive patterns and trends that can help predict Court outcomes. And this phenomenon is made possible by the capacity for technological «manipulation» with humanly inaccessible speed of huge amounts of big data that are extremely heterogeneous and cannot be stored by humans.

What used to appear as a mere Enlightenment ideal now seems to have become a «historical reality»[4]. However, the question, even before whether this is practically feasible, is whether it is desirable. Indeed, it is undeniable that predictive justice raises questions on at least two levels. Firstly, in terms of its legitimacy and legality and, secondly, in terms of its quality and reliability. Specifically, developing the use of systems of this kind presupposes the acceptance of the fact that a machine learning algorithm – by means of natural language processing and deep learning techniques – analyses an indefinable a priori number of big data (represented, for example, by regulatory provisions and case law precedents). This, in order to assess, on the basis of a statistical calculation and inductive-probabilistic reasoning, the validity or unfoundedness of a given legal solution in relation to a specific question of fact, or of disputed law. Replacing, therefore, the judge-human with a judge-robot.

2. What's good

Before proceeding with the analysis of what, in the opinion of the majority of scholars, are the reasons that hinder – not only nowadays, but tend to definitively – the adoption of a full predictive justice system, it is appropriate to take note of what, at the same time, are the advantages that the legal system could draw (and which it timidly draws today) from the use of Artificial Intelligence. It should be premised that, in historical reality, predictivity applied to justice is not a new theme, as some of the writings of Poisson[5] and Condorcet[6] note.

[4] B. Occhiuzzi, *Algoritmi predittivi: alcune premesse metodologiche*, in «Dir. pen. cont. – Riv. trim.», II (2019), p. 393.

[5] S.D. Poisson, *Recherches sur la probabilité des jugements en matière criminelle et en matière civile,* Bachelier, Parigi, 1837.

[6] M. De Condorcet, *Essai sur application de l'analyse à la probabilité des décisionsrendues à la pluralité des voix*, Imprimerie Royale, Parigi, 1785.

In more recent times, the issue has been addressed in its most diverse aspects in the preponderance of national legal systems. One thinks, merely by way of example, of Loevinger's studies on so-called jurisprudence[7] and Reed C. Lawlor's investigation into the usefulness of the use of computers and algorithms in scientific legal research and trial[8]. Or to the international conferences on the operation of so-called expert systems[9] held in Bologna and Florence respectively, and, most recently, to the introduction in France of Law 2016-1321 (so-called «Legge per la Repubblica digitale») and the establishment of the online platform Prédictice[10]. In Italy, too, there are the first examples of AI experimentation in the justice sector, such as Prodigit, a project in the tax field that aims to create systems capable of analyzing laws, judgements and doctrinal contributions to predict legal decisions. Of no less significance is the initiative promoted in 2021 by the Court of Appeal, the Court of Brescia and the University of Brescia to provide users with all elements of «Labour law» and «Company law», in order to discourage reckless lawsuits and promote the transparency of decisions. Or the collaboration agreement between the CED of the Court of Cassation and the Higher University School of Pavia (IUSS).

Indeed, it is undeniable that a system that can take advantage of the fruits of technological evolution gains (albeit with the clarifications that follow) in terms of effectiveness, efficiency, transparency, speed, and the litigation's deflation. On one hand, legal practitioners could have wider and, above all, faster access to information. They would benefit from a better knowledge and understanding of the regulatory, jurisprudential and factual situation inherent in the case. On the other hand, since technological tools could provide hypotheses for resolving the dispute or individual procedural or substantive aspects, this would certainly increase the calculability or, in other words, the predictability of decisions. This, moreover, represents a more specific declination of the principle of legal certainty. It would be no less beneficial in terms of time, since the use of machines would considerably reduce the «length of proceedings» so as to guarantee a totally new (and by most people hoped for) speed of judgements. At the same time, AI could suggest to lawyers themselves as well as to citizens to avoid judicial solutions in favor of new out-of-court dispute resolution opportunities such as negotiation or mediation. In addition,

[7] L. Evinger, *Jurimetrics: The Next Step Forward*, in «Minnesota Law Review», 1949.

[8] C. Lawlor, *What Computers Can Do: Analysis and Prediction of Judicial Decisions*, in «American Bar Association Journal», 1963.

[9] L. Lombardi Vallauri, *Verso un sistema esperto giuridico integrale*, in «Jus», II (1996), p.157 ss.

[10] It is used to calculate the probability of a case being settled and the amount of damages obtained in similar litigation, on a database consisting of the judgments of the Courts of Appeal and the Supreme Court (www.predictice.com/fr).

systems and mechanisms for calculating compensation in civil matters (e.g. compensation for bodily harm, severance pay) could be set up (or, as they already exist in some national laws, increased). Last but not least, decisions would be based on a more «objective, neutral, complete and accurate examination» of the documents and information contained therein and would also be less exposed to the risks of «noise, i.e. judicial subjectivism»[11]. It seems appropriate to dwell for a moment on the advantage that the system would derive in terms of transparency from the public dissemination of judgments and court orders on the web portals of the judicial authorities. While this, indeed, plays a fundamental role in guaranteeing the free search of all decisions, their full disclosure, however, exposes the natural persons mentioned in the decisions to individual risks of violation of rights (including the right to privacy) and fundamental freedoms. Part of the doctrine has attempted to obviate this risk and, therefore, to reconcile the necessary respect for the rights of individuals with the requirements of the circulation of information in a democratic state. The solution can be found in the so-called right to anonymization of judgments[12], which implies the removal, ex officio or on request of a party, of personal data contained in the text of a judgment or Court order before it is released to the public. Once 'cleansed' of any personal reference, the documents could be fed into a predictive justice system that provides an assessment of the likelihood of winning a dispute or to automatically derive case law maxims.

3. The other side of the coin

Having acknowledged the countless and undeniable advantages that the justice system would be able to draw from the use of predictive justice systems, it is necessary to duly highlight those that, at the same time, and at the moment certainly of greater weight and relevance, are the risks, limitations and illusions connected to the substitution of the human figure of the judge with a machine. As we shall see in more detail, such and so many are the obstacles and dangers, that they outweigh the benefits. Before proceeding, one can summarize this thought by evoking the emblematic expression coined by G. Calabresi to describe progress. He speaks of it by referring to a «gift of the Evil spirit», which offers itself to humanity to make its life more and more comfortable, but which in time proves to be anything but free, impo-

[11] See, on the subject, D. Kahneman, O. Sibony, R. Cass Sunstein, *Rumore*, UTET, Torino, 2021, p. 23.

[12] The right to anonymization of judgments is governed by Regulation (EU) 2016/679 and Article 52 of the Privacy Code.

sing heavy costs of which the correlation with the benefits is not immediately perceivable[13]. From this perspective, since predictive justice is an instrument that aims to defeat and oust the judge in flesh and bones, it must necessarily be scrutinized by analyzing its compatibility with the legal frameworks of International, European and National derivation. Indeed, an improper use of this technology could pose a danger not only to human rights and freedoms, but also to the democratic order at all.

A first insurmountable limit to a complete transfer of decision-making power from man to machine is to be found in the specificity of individual cases. In other words, to borrow an expression used by N. Irti, there is no repetitiveness in law[14]. Every case of everyday life differs from another, because subjects, causes, feelings, passions and other infinite aspects that have contributed to the emergence of that relationship or event are different. From this it follows that, if the machine draws its decision from the analysis of past decisions on similar cases, it will never be able to find an answer, since it does not have and does not exist «a (fully) analogous case» from which to translate it.

Closely related to the limitation just mentioned, as a further and fundamental obstacle to the «mechanization» of the trial, it is appropriate to dwell on the role of the judge. Indeed, the activity performed by him is never ends up in a merely mechanical operation of applying the rule of law. However, one must be careful not to run into a misunderstanding. In fact, what has just been said is not intended to attribute to the decider a legislative function that, according to Montesquieu's tripartition, belongs to Parliament. But, at most, it is intended to make people aware of an indisputable reality, namely that Jurisprudence, even if it does not create it, «makes the law»[15]. Above all, this if one considers modern societies, which are in constant flux, and whose dynamics are often too fluid and lively to be foreseen and adequately regulated by predetermined and specific rules, also because of the long technical time it takes to pass a law. A judge and a living law that do not oppose but are well reconciled with the physiological presence of gaps in legal protection.

The identification of the actual content of the rule to be applied to the concrete case inevitably passes through an interpretative activity to which, first of all, the judges of the merits take part, not limiting themselves to select and apply the rules, but going further by providing a reading in accordance with Constitution, European Union Law and Conventional Law. But, it is not all. In addition to identify the rule to be applied among a multitude of sources that

[13] G. Calabresi, *Il dono dello spirito maligno: gli ideali, le convinzioni, i modi di pensare nei loro rapporti col diritto,* Giuffrè, Milano, 1996.

[14] N. Irti, *Un diritto incalcolabile,* Giappichelli, Torino, 2015.

[15] G. Gorla, *Diritto comparato e diritto comune europeo,* Giuffrè, Milano, 1981, p. 263.

do not fit together in a perfectly coherent manner (the legal theorist Herbert L.A. Hart described the law as an «open fabric»[16]) and to interprete it, Jurisprudence must also, for decision-making purposes, select and evaluate a priori indefinable multitude of relevant elements and facts.

The latter, which can be qualified as profiles of the specificity of the case, as essential elements of the individual disputed case, necessarily influence and calibrate the judicial decision to such an extent that a machine capable of recognizing them a priori is not conceivable. On the contrary, they are necessarily harbingers of the still typically human activity of distinguishing, which in turn is balanced by an obligation of reasonable and exhaustive reasoning. Thus, if deciding is not the mechanical application of rules, a decision-making process based on statistical generalization of previous rulings would guarantee what could be called the «decision of the average judge», but certainly not the right one. In this sense, emblematic is the conviction expressed by Alan Turing, who stated that a machine would never be able, among other things, to «*avere iniziativa, avere senso dell'humour, distinguere il bene dal male, commettere errori, innamorarsi, gustare le fragole con la panna*»[17]. Or the words of Piero Calamandrei who wrote «*(n)oi non sappiamo più che farci dei giudici di Montesquieu, 'êtres inanimés' fatti di pura logica. Vogliamo dei giudici con l'anima, giudici engagés, che sappiano portare con vigile impegno umano il grande peso di questa immane responsabilità che è il rendere giustizia*»[18]. Therefore, a mechanical decision-maker hides the danger not only of a trivialization of the decisions, but also of losing the ineradicable emotional and valuable dimension of the judgement that, resting on the «innumerable nuances» of human ideas and emotions, has «paradoxically its strong point in imperfection»[19].

Moreover, a system founded on the idea of the Automaton Judge, who decides on the basis of the analysis of judicial precedents (in addition, of course, to the data of the case and the rules of law), inevitably heads towards a standardization of judicial decisions, thus preventing the system itself from progressing and evolving in line with the change of the context of reference[20].

[16] H.L.A. Hart, *Il concetto di diritto*, Einaudi, Torino 2002 (or. ed.: *The Concept of Law*, Clarendon Press, Oxford 1961).

[17] A. Turing, *Macchine calcolatrici e intelligenza* (1950), in *Intelligenza meccanica*, a cura di G. LOLLI, Bollati Boringhieri, Torino 1994, p. 139.

[18] P. Calamandrei, *Opere giuridiche. Volume I. Problemi generali del diritto e del processo,* Roma, RomaTrE-Press 2019, p. 650.

[19] G. Zaccaria., *Figure del giudicare: calcolabilità, precedenti, decisione robotica*, in «Riv. dir. civ.», II (2020), p. 2.

[20] Very interesting on the subject, and to which it refers, is the comparison between law and the Galton Machine developed by da C. Barbaro, *Uso dell'intelligenza artificiale nei sistemi giudiziari: verso la definizione di principi etici condivisi a livello europeo?*, in «Questione Giustizia», IV (2018).

Such failure to change would bring with it an unjustified injury to the right of individuals to obtain a fair decision, which constitutes the expression of the primary function of law, that is to provide answers to human needs by regulating relations in a manner corresponding to the social needs of the particular historical moment. Therefore, it would be legal immobilism due to an imposition of a – even mechanical – constraint of automatic conformation. And mind you, Common Law systems themselves tend to stem this type of risk, even though the system historically rests on the rule of *stare decisis,* stimulating a figurative dialogue between the judge of the present case and those of the past case with the intention of verifying the applicability of the same decisional principle also to the second case decided.

At the same time, the acceptance of a pure predictive system would not only compromise – with a view to crystallization – the Jurisprudence as a whole but would also have negative effects on the individual magistrates in terms of impartiality and de-responsibility. These, indeed, would rely on the 'answers' provided by the algorithm as much out of superficiality as for reasons of «defensive justice». Specifically, on the one hand, the algorithmic solution would develop an 'accommodating' mechanism insofar as the lazy judge would no longer go beyond the simple result provided by the machine, accepting it in a tendentially passive manner and without proceeding to an analysis of the question in light of the particularities of the individual case. Even defined by some as the «performative» effect of technological prediction, it no longer has an indicative value, but an almost prescriptive one for the human decision-maker. At the same time, on the other hand, ruling against an algorithmic prediction implies the assumption of a risk and, consequently, greater responsibility and pressure. It does not seem unrealistic to imagine, as appropriately reported by some doctrine, «that judges may be reluctant to shoulder this burden, particularly in systems where their mandate is not permanent, but subject to popular vote, or in those where their personal responsibility (disciplinary, civil or even criminal) may be enforced, especially if the guarantees provided by law on disciplinary matters are insufficient»[21]. And, perhaps, even worse – on an ideal level – would be the practices of scoring, or social «scoring», i.e. the monitoring and ranking of judges or lawyers based on decisions taken or conduct.

However, the risks associated with technological development in the judicial field do not end with those mentioned above, but concern so many more aspects. In fact, the data put into the machine can produce negative effects on the decision and, in general, a deterioration of the justice system. Indeed,

[21] C. Barbaro, *Uso dell'intelligenza,* cit.

they not only contribute to producing that phenomenon of the crystallization of law referred to earlier (being grammatically past participles nounized), but, even more seriously, the selection of the same data and their correlations can consecrate possible preconceptions[22]. Indeed, the entire pathway from data collection to results is affected by individual discretion, as it is necessary not only to choose the type of elements to be filed, but also to establish their requirements, their minimum number to consider them qualitatively and quantitatively sufficient to be significant, and so on. The whole group of these latter activities is carried out, in turn, by a human subject who, therefore, suggests to the machine to grasp what he wants it to grasp, so much so that part of the doctrine has affirmed that the algorithm «is ontologically conditioned by the value system and intentions of those who commission its creation and/or those who create it»[23]. In this regard, the case concerning the famous Loomis case, decided by the Wisconsin Supreme Court about assessing the danger of recidivism for the purpose of quantifying the sentence for a person already found guilty, was much discussed. On that occasion, it was shown, as a bias of the machine itself, that the commitment and full remittance of the decision to the COMPAS system revealed the existence of mainly racial discrimination. In addition to all that has been said in relation to the relationship between machine and owner in terms of discretion and subjectivity, there is also another aspect that should be emphasized. Precisely, being technological instruments that are produced within private companies, they are protected by privative rights or industrial secrecy, with all that this entails in terms of the risk of opacity in the selection of data and calculation and evaluation algorithms.

Lastly, a further victim of the Automaton Judge would be the legal idea of the «judgement of equals», expressly sanctioned by the Magna Charta Libertatum since 1215 and, as of today, enshrined in Article 111 of the Constitution. There would no longer be two subjects placed on the same level and «competing» in respect of the adversarial process to obtain what they claimed before a third and impartial subject. On the contrary, there would be a machine that, not giving the possibility to fight on equal terms, would decide independently and above the contenders.

All of the above would place predictive justice in a position of clear conflict with the foundations of our judicial system. There would be a tacit abdica-

[22] A. Garapon, J. Lassegue, *Justice digitale. Révolution graphique et rupture anthropologique,* PUF, Parigi, 2018, p. 257 ss.; L. Maldonato, *Algoritmi predittivi e discrezionalità del giudice: una nuova sfida per la giustizia penale,* in «Dir. pen. cont. – Riv. trim.», II (2019), p. 407; J. Nieva-Fenoll, *Intelligenza artificiale e processo* (2018), Giappichelli, Torino 2019, p. 65.

[23] S. Signorato, *Giustizia penale e intelligenza artificiale. Considerazioni in tema di algoritmo predittivo*, in «Riv. dir. proc.», II (2020), p. 614.

tion of the role of the judge, in clear conflict with the characteristics of naturalness and pre-constitutionality (Art. 25 paragraph 1 Const.), the nullification of the duty, established by Art. 111 paragraph 6 Const., of a motivation controllable by the people in whose name «justice is administered», the denial of jurisdictional subjection «only to the law» (Art. 101 Const.) and the frustration of the adversarial process[24].

It would be appropriate, therefore, to adopt a somewhat aseptic view of the role that predictive justice may be called upon to play in the judicial sphere. And both supranational and national legislators are moving in this direction, each working to outline regulatory and para-regulatory coordinates capable of guiding and regulating the behavior of operators, including legal ones, in the Artificial Intelligence sector. Indeed, the European Commission on the Efficiency of Justice (Cepej) has conducted an in-depth study on these issues and drafted a «Carta etica europea sull'uso dell'Intelligenza Artificiale nei sistemi giudiziari»[25]. which sets out substantive and methodological principles for the production and implementation of algorithmic tools, also applicable to the analysis and processing of judicial decisions.

Among the guiding set out criteria are the respect for human rights and the non-discrimination by AI applications through full compliance with the rights guaranteed by the European Convention on Human Rights and the Council of Europe Convention on the Protection of Personal Data, from the design stage to the practical application. Of no less importance is the principle of transparency, i.e. the need to make data processing techniques accessible and understandable, as well as to authorize external audits by independent parties in order to identify any distortions. The processing of information carried out by machine learning, indeed, should be carried out on the basis of knowable data (about which the subjects must be informed in a clear and comprehensible language) and which are concrete expressions of the different realities on which the judge is called to rule. At the same time, it is essential to guarantee the possibility for the latter to deviate from the algorithmic results in view of the specificities of the particular case.

A primary role within the discourse that is developing, however, is the discipline adopted about the relationship between algorithm and judge. Indeed, even in the face of the well-known Loomis case, European bodies believe

[24] G. Fanelli, *L'impiego dell'intelligenza artificiale nei processi decisori del giudice, tra la disciplina europea e quella del processo civile, in Il diritto nell'era digitale. Persona, Mercato, Amministrazione, Giustizia,* a cura di R. Giordano *et al*., Giuffrè, Milano 2022, p. 993; M. Luciani, *La decisione giudiziaria robotica*, in *Decisione robotica*, a cura di A. Carleo, Il Mulino, Bologna 2019, p. 86.

[25] The first European instrument on this subject was approved by the Council of Europe during the CEPEJ plenary session of 3-4 December 2018 (available at https://rm.coe.int) referring in turn to a study conducted by University College London (UCL), which dealt with 584 machine learning decisions of the European Court of Human Rights with a success rate of 79%.

that for qualified «high-risk» systems (which would include those of predictive justice) there must necessarily be some form of human supervision, and they also warn of the risk of over-reliance by humans in the development of Artificial Intelligence (so-called automation bias). The Charter of Ethics, although, as initially stated, it is not a binding legal act for states, has had a pervasive effect at the global level, having been a clear source of inspiration from many quarters. By way of example, we recall the adoption of the US AI Bill of Rights or, at the European level, the Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law, adopted by the Council of Europe on 17 May 2024, the aim of which is to ensure that, in the international context, the activities included in the entire lifecycle of AI systems respect these values[26]. Or the very recent EU Reg. 2024/1689 of 13 June 2024 on Artificial Intelligence, also known as the AI Act. However, it should be pointed out that even in the earlier era, the European legislator felt the need to curb the risks inherent in the use of algorithmic systems. Examples of this approach can be found in: Art. 22 Reg. No 2016/679/EU, Art. 11 Di. No 2016/680/EU and Art. 8 Legislative Decree No 51 of 2018, where the so-called «principle of non-exclusivity» of the algorithmic decision is enshrined in the privacy law. They exclude, that is, decisions supported solely by automated processing, including profiling, since the exceptions – which nevertheless require adequate safeguards for individual rights, freedoms and legitimate interests – must also provide for human intervention in the automated decision-making process relating to natural persons so as not to harm the dignity.

For its part, Italy was one of the first European countries to act in compliance with supranational guidelines and on 20th May 2024 submitted for parliamentary approval the bill no. 1146, entitled «Disposizioni e delega al Governo in materia di Intelligenza Artificiale». Clearly, the will of the national legislature to promote a responsible use of Artificial Intelligence. To this[27] end, it seems essential to put man and respect for fundamental rights at the center of the system, as well as to guarantee a) the fair and correct use of algorithms; b) the quality, security and protection of the data used and the processes with which they are processed and processed by AI systems; c) digital sustainability. As regards specifically the field of justice, they note Articles 14 and 15, which, in summary, reaffirm the concept of human in command. This expression means that it is legitimate to use and benefit from the advantages arising from the use of algorithmic systems, but this is only for the purpose of simplifying and assisting the figure of the natural person judge, in charge of which the responsibility and the duty of decision persists. It is the same art.

[26] The document is available at the Council of Europe's institutional address www.co e.int.

[27] The d.d.l. A.S. n. 1146 can be read on the Senate institutional website www.senato.it.

14, in fact, to state that «*Artificial Intelligence systems are used exclusively for the organization and simplification of judicial work as well as for jurisprudence and doctrinal research*». The Ministry of Justice regulates the use of Artificial Intelligence systems by ordinary judicial offices. For the other jurisdictions the use is regulated in accordance with the respective laws. It is always reserved to the magistrate the decision on the interpretation of the law, on the evaluation of facts and evidence and on the adoption of any measure". Already there are multiple applications of AI on the market that could be used in the pursuit of the purposes prescribed by the bill. To name a few, consider the implementation of the potentialities of some legal databases – specifically, the Public Merit Database (BDP) and the Merit Confidential Database (BDR), accessible only to magistrates – using generative Artificial Intelligence systems. Or even the development of algorithmic tools that can identify, through the computational analysis of data, any organizational gaps – so as to promote a better allocation of resources – or AI tools for the translation of documents, for the analysis and revision of documents.

Technological devices to support the work of the judge, by reducing the area of judgment, ensure greater predictability of decisions, through the use of algorithms capable of analyzing large amounts of legal data (including legislative texts, judicial precedents and decisions of courts, courts of merit, the Court of Legitimacy and supranational ones) and offer to the same judge tools to strengthen the rational and decision-making faculties. To reduce, it should be noted, does not mean to eliminate the necessary control and the consequent responsibility of the judge-man. In the future, therefore, artificial justice programs (including predictive ones) – if respected and guaranteed certain conditions – will become an indispensable aid for the work of the judge and perform a supporting function to the same. With subsequent meaningful human control, we would have the desired integration between algorithm and decision-making in light of fundamental principles of our law. AI thus becomes a tool, not an end.

4. Beyond good and evil

All this inevitably leads to admitting, not without a note of subtle relief, that the performance of Artificial Intelligence applied to jurisprudence is – at least to date – comparatively weaker and more imperfect than in other fields. Despite this, it would be an important mistake not to exploit its great potential to support and improve the hard work of the judiciary. Indeed, «Predictive justice» is only one of the many applications of Artificial Intelligence, but there are other, very interesting, ones being trailled in some European countries.

If you want to mention a few, think of the use that is being made in Austria for the quick reading, classification and attribution of acts, appeals and documents to the relevant chancery sections; or the use of chatbots (literally: conversation with a robot) to direct the citizen towards an alternative solution of the dispute, as in Latvia.

Taking note of all that has been said in terms of the potentials and the risks of a new use of AI tools in the field of law, it is considered to conclude with a consideration that, even before it is conclusive, could and should be understood as a preliminary and fundamental aspect for an ordinal evolution in this direction. The challenges posed by technological development can only be tackled more effectively through a fusion of knowledge. In this regard, as pointed out by more authoritative jurists and scholars, the creation of multi-disciplinary teams composed of magistrates, researchers in social sciences and computer science would benefit at the stage as much of elaboration as of piloting and implementing the solutions proposed by science.

BIBLIOGRAPHY

Barbaro C., *Uso dell'intelligenza artificiale nei sistemi giudiziari: verso la definizione di principi etici condivisi a livello europeo?,* in «Questione Giustizia», IV 2018.

Battelli E., *Giustizia predittiva decisione robotica e ruolo del giudice*, in «Giustizia Civile», II (2020), p. 281-319.

Calabresi G., *Il dono dello spirito maligno: gli ideali, le convinzioni, i modi di pensare nei loro rapporti col diritto*, Giuffrè, Milano 1996.

Calamandrei P., *Opere giuridiche. Volume I. Problemi generali del diritto e del processo*, Roma-TrE-Press, Roma 2019, p. 650.

De Condorcet M., *Essai sur application de l'analyse à la probabilité des décisionsrendues à la pluralité des voix*, Imprimerie Royale, Parigi, 1785.

Fanelli G., *L'impiego dell'intelligenza artificiale nei processi decisori del giudice, tra la disciplina europea e quella del processo civile*, in *Il diritto nell'era digitale. Persona, Mercato, Amministrazione, Giustizia*, a cura di GIORDANO R. *et al*., Giuffrè, Milano 2022, p. 993.

Galimberti A., *Psiche e techne. L'uomo nell'età della tecnica*, Feltrinelli, Milano 2016.

Garapon A., Lassegue J., *Justice digitale. Révolution graphique et rupture anthropologique*, PUF, Parigi, 2018.

Gorla G., *Diritto comparato e diritto comune europeo*, Giuffrè, Milano 1981.

Hart H.L.A., *Il concetto di diritto*, Einaudi, Turino 2002 (ed. or.: *The Concept of Law, Clarendon Press*, Oxford 1961).

Irti N., *Un diritto incalcolabile*, Giappichelli, Torino, 2015.

Kahneman D., Sibony O., Cass Sunstein R., *Rumore*, UTET, Torino 2021.

Lawlor C., *What Computers Can Do: Analysis and Prediction of Judicial Decisions*, in «American Bar Association Journal», (1963).

Loevinger L., *Jurimetrics: The Next Step Forward*,in «Minnesota Law Review», (1949).

Lombardi Vallauri L., *Verso un sistema esperto giuridico integrale*, in «Jus», 1996, p.157 ss.

Luciani M., *La decisione giudiziaria robotica*, in *Decisione robotica*, a cura di CARLEO A., Il Mulino, Bologna 2019, p. 86.

Maldonato L., *Algoritmi predittivi e discrezionalità del giudice: una nuova sfida per la giustizia penale*, in «Dir. pen. cont. — Riv. trim.», II (2019), p. 407.

Nieva-Fenoll J., I*ntelligenza artificiale e processo*, Giappichelli, Torino 2019, p. 65.

Occhiuzzi B., *Algoritmi predittivi: alcune premesse metodologiche,* in «Dir. pen. cont. – Riv. trim.», II (2019), p. 393.

Poisson S.D., *Recherches sur la probabilité des jugements en matière criminelle et en matière civile*, Bachelier, Parigi, 1837.

Satta S., *Il mistero del processo*, in «Riv. dir. proc.», I (1949).

Signorato S., *Giustizia penale e intelligenza artificiale. Considerazioni in tema di algoritmo predittivo*, in «Riv. dir. proc.», (2020), p. 614.

Turing A., *Macchine calcolatrici e intelligenza (1950)*, in *Intelligenza meccanica*, a cura di Lolli G., Bollati Boringhieri, Torino 1994, p. 139.

Zaccaria G., *Figure del giudicare: calcolabilità, precedenti, decisione robotica*, in «Riv. dir. civ.», II (2020), p. 2.

SIMONE SACCOMANI*

The accountability and liability paradigms applied to AI: the European framework§

Abstract

The article, after a preliminary theoretical analysis around the functional paradigm shift on liability, and the damage caused by intelligent products and AI systems, systematically reflects on the regulatory proposals at the Euro-EU level. After an overview of the doctrinal debate around accountability and liability with respect to AI, a reconstruction of these principles and their relative functions and main areas of application will be proposed, with regard to the AI Liability Directive and the new Product Liability Directive, in order to examine the advisability of implementing a strong coordination with civil liability. Subsequently, some critical insights from the liability analysis will be provided to see how well this approach fits into the emerging legal framework. In the conclusions, some criticisms and suggestions for possible improvements of the existing disciplines will be provided.

1. Introduction

Artificial Intelligence (AI) represents a technological phenomenon sui generis compared to the symbolic technologies of previous industrial revolutions[1]. Its peculiarity essentially consists in not actually being a "new techno-

* PhD candidate Università della Svizzera italiana; Visiting Professor Istituto Universitario Progetto Uomo

§ This contribution develops the presentation at the conference 'Cybersecurity e prodotti: sfide e responsabilità nelle società interconnesse',29 November 2024, Italian Supreme Court.

[1] M. Callier, H. Callier, *Blame it on the machine: socio-legal analysis of liability in an ai world,* in «Washington Journal of Law, Technology&Arts»,XIV, I (2018), pp. 49-76.

logy", but in representing an approach that transversally invests numerous technological sectors, with the related implications in terms of Cybersecurity[2]. "Modern" AI differs from "classical" AI for the mixture of machine learning techniques and different technological phenomena, ex multis the development of the Internet, computers and Big Data, which makes it possible to associate current AI systems with certain peculiar and unprecedented characteristics compared to traditional products[3]. AI systems can autonomously make decisions, learn from experience and evolve over their operational dimension, with a margin of unpredictability, which risks causing humans to lose control and supervision over algorithmic decision-making[4].

At the methodological level, the current national and Euro-unitary substantive law will be examined with reference to the most recent regulations, carrying out comparative in-depth studies.

In the current regulation of AI, the most recent supranational standardization interventions, especially the AI Act, correspond to significant margins of application of existing national law, particularly in the field of Cybersecurity and civil liability[5]. It is therefore necessary to examine the most current reflections of the doctrine, and the Euro-unitary regulatory proposals relating to the safety of smart products and the corresponding civil liability, to assess the impact that these disciplines could have on the balance of the current liability system[6].

Recent doctrine is in favour of a renewed centrality of the preventive function of liability with respect to its classic compensatory function, favouring a regulatory evolution of AI in which the principle of accountability assumes primary importance in the general regulatory framework in which civil liability becomes a component of the complex governance structure inspired by the proceduralization of certain ethical and legal purposes within the technical-organizational mechanisms, and the preventive function tends to "invade" the field of civil liability[7]. In the regulation of contemporary technological innovation, the concept of liability increasingly refers to the dimension

[2] T. De Mari Casareto Dal Verme, *Intelligenza artificiale e responsabilità civile. Uno studio sui criteri di imputazione,* Edizioni Scientifiche Italiane, Napoli, 2024, p. 320.

[3] G. Di Rosa, *Quali regole per i sistemi automatizzati intelligenti*, in «Rivista di diritto civile», V (2021), pp. 823-853.

[4] A. Etzioni, O. Etzioni, *Keeping ai legal,* in «Vanderbilt Journal of Entertainment & Technology Law», XIX, I (2016), pp. 133-146.

[5] G. Capilli, *Responsabilità civile e Robot*, in «La nuova giurisprudenza civile commentata», III (2019), p. 628.

[6] V. Di Gregorio, *Intelligenza artificiale e responsabilità civile: quale paradigma per le nuove tecnologie?,* in «Danno e responsabilità», I (2022), p. 61.

[7] «*Algorithmic accountability does not only require the examination of algorithms or the code as such but also an examination of how algorithms are deployed within different areas and what*

of accountability, where civil liability, while maintaining conceptual autonomy, becomes a component of the new general framework represented by «a governance structure aimed at certain ends and governed by precise legal and ethical constraints preordained and oriented, precisely, to establish such technical or organizational mechanisms»[8].

2. The principles of accountability and liability applied to AI

The principle of accountability[9] originates from the Anglo-Saxon legal context, where the term has taken on a complex scope and is often used in public law and finance as one of the key principles[10]. This discipline requires the decision-maker to adopt, in advance and from the beginning of his activity, technical and organizational measures to eliminate the potential negative effects of his decisions[11]; moreover, it implies the obligation to justify one's decisions before the addressees of those decisions and before an external authority[12]. Accountability means both "responsibility" and "proof of responsibility" to the extent that the decision-maker must demonstrate that he or she has prepared the measures envisaged at the request of the data subject or a third party, who may impose sanctions where violations are detected, even in the absence of actual damage suffered by the data subjects[13]. Raffaele Marcello defines accountability as a complex concept, which is based on the combination of autonomy and responsibility of the governing bodies of companies towards stakeholders, whose breadth of content calls into question a plurality of multidisciplinary, socio-economic and legal skills[14].

the tasks are that they perform» A. Vedder; L. Naudts, *Accountability for the use of algorithms in a big data environment,* in «International Review of Law, Computers & Technology», 31, II, (2017), pp. 206-224.

[8] G. Comandè, *Lettera sulla responsabilità (civile) e l'autonomia (individuale),* in «Danno e responsabilità», VI (2022), p. 667.

[9] R. Mulgan, *Holding Power to Account: Accountability in Modern Democracies,* Palgrave Macmillan, London, 2003, passim.

[10] L. Polverari, *Alcune riflessioni sul l'accountability della politica di coesione europea, sul trade-off tra diversi tipi di accountability e sul trade-off tra accountability ed efficacia,* in «Rivista giuridica del Mezzogiorno», IV (2011), pp. 1093-1108.

[11] E. Chiti, *L'accountability delle reti di autorità amministrative dell'Unione europea,* in «Rivista italiana di diritto pubblico comunitario», XXII, I (2012), pp .29-84.

[12] M. Bovens, *Two Concepts of Accountability: Accountability as a Virtue and as a Mechanism*, in «West European Politics», XXXIII, V (2010), pp. 946-967.

[13] P. Perlingieri, S. Giova, I. Prisco (a cura di), *Rapporti civilistici e intelligenze artificiali: attività e responsabilità. Atti del 15° Convegno Nazionale 14-15-16 maggio 2020*, Edizioni Scientifiche Italiane, Napoli 2020, p. 42.

[14] R. Marcello, *Accountability: una prospettiva politologica*, in «Enti non profit», VII (2007), pp.457-461.

Accountability represents the civil liability[15] related to a given role and indicates the claim that decision-makers [16] are called to account for their choices and assume responsibility for any negative effects of them, including through mechanisms that involve the reversal of the burden of proof[17].

Despite the complexity that it inherently encompasses, two peculiar characteristics of accountability can be identified[18].

In primis, its application context is purely vertical, in the sense that accountability translates into expectations and control tools that are expressed in the relationship between the recipients of unilateral decisions, decision-makers and public supervisory authorities[19]. Consequently, when the principle is expressed explicitly in provisions of positive law, these assume a predominantly public nature.

Consequently, when the principle is made explicit in rules of positive law, these take on a predominantly public nature. The second characteristic of accountability is that it performs a primary function of prevention, since it is oriented towards the ex-ante responsibility of the decision-maker regarding the countermeasures to be adopted from the first moment in which he is called upon to provide the means by which to achieve the expected results, although the remedies also include the imposition of ex-post sanctions that fully implement the preventive-deterrent function of accountability[20].

Therefore, accountability differs from liability because the former refers to a moment prior to the possible occurrence of detrimental events caused by the decisions taken, while the latter deals with transmitting ex post the cost of the damage caused to the legal sphere of the victim to that of a third party held liable by the regulations in force[21].

The most recent hermeneutical proposals relating to instances developed by recent doctrine are oriented towards a favourable approach

[15] M. Barcellona, *La responsabilità civile,* a cura di S. Mazzamuto, *Trattato di Diritto Privato,* Giappichelli, Torino 2021, pp. 4-8.

[16] A. Cassatella, *La responsabilità funzionale nell'amministrare. Termini e questioni,* in «Diritto amministrativo», III (2018), pp. 677-746.

[17] G. Comandè, *Responsabilità ed accountability nell'era dell'Intelligenza Artificiale*, a cura di F. Di Ciommo, O. Troiano, *Giurisprudenza e autorità indipendenti nell'epoca del diritto liquido. Studi in onore di Roberto Pardolesi,* Piacenza, La Tribuna, 2018, pp. 1001-1013.

[18] G. Teubner, *Soggetti giuridici digitali? Sullo status privatistico degli agenti software autonomi*, Edizioni Scientifiche Italiane, Napoli, 2019, pp. 27-30.

[19] M. Bussani, *Le funzioni delle funzioni della responsabilità civile,* in «Rivista di diritto civile», II (2022), pp.264-306.

[20] T. De Mari Casareto Dal Verme, cit., p.328.

[21] C. Salvi, *Quali funzioni per la responsabilità civile?*, in «Danno e responsabilità», II (2022), pp. 137-140.

to the regulation of AI inspired by the principle of accountability and its essential preventive function[22].

Part of the doctrine points out that mere liability rules do not guarantee suitable incentives for prevention, as they would be inadequate in dividing economic responsibility among the entities located in the value chain of AI systems[23]. According to this approach, the plurality of entities involved in the production of AI systems would pose the need to allocate the financial risk exactly among those operators, a need that would not be fully satisfied by the provision of joint and multiple liability, which makes it possible to simplify the procedural position of the injured party, but does not correctly allocate the costs of prevention[24].

Another part of the doctrine, also in favour of the accountability approach, emphasizes the need to make the subject advantaged by the application of AI responsible, who should bear the burden of taking the necessary measures to avoid the risk and demonstrating the way in which he has worked to avoid it[25].

The preference for accountability in the standardization of AI systems is legitimized by the intrinsic complexity of the operation of intelligent algorithms and, above all, by the multiplicity of economic entities that pertain to the product value chain, so much so that it is difficult to attribute responsibility for malfunctioning only to a specific operator[26]. The overlapping of roles and competences would make it particularly difficult to trace the progression of causal links and, therefore, to identify the responsible party[27].

The principle of accountability would allow the sharing of responsibilities between these entities, encouraging economic operators to adopt preven-

[22] F. Faini, *Il diritto nella tecnica: tecnologie emergenti e nuove forme di regolazione,* in «Federalismi.it», XVI (2020), pp. 79-117.

[23] L. Coppini, *Robotica e intelligenza artificiale: questioni di responsabilità civile,* in «Politica del diritto», 4 (2018), pp. 713-740; G. Comandè, *Intelligenza artificiale e responsabilità tra liability e accountability. Il carattere trasformativo dell'IA e il problema della responsabilità,* in «Analisi Giuridica dell'Economia», I(2019), pp. 169-188; G. Comandè, *Multilayered (Accountable) Liability for Artificial Intelligence,* a cura di S. Lohsse, R. Schulze, D. Staudenmayer, *Liability for Artificial Intelligence and the Internet of Things,* Baden-Baden, Nomos, 2019, pp. 165-184.

[24] D.C. Vladeck, *Machines Without Principals: Liability Rules and Artificial Intelligence,* in «Washington Law Review», LXXXIX (2014), pp. 117-149.

[25] M.C. Gaeta, *Liability rules and self-driving cars. The evolution of tort law in the light of new technologies*, Edizioni Scientifiche Italiane, Napoli 2019, passim; G. Finocchiaro, *Intelligenza artificiale e protezione dei dati personali*, in «Giurisprudenza Italiana», VII (2019), pp. 1670-1677.

[26] L. Gatt, I. Caggiano, M.C. Gaeta, *Italian Tort Law and Self-Driving Cars: State of art and Open Issues,* a cura di B.H. Oppermann, J. Stender-Vorwacs, *Autonomes Fahren. Technische Grundlagen, Rechtsprobleme, Rechstsfolgen,* München, Beck, 2020, pp. 239-255.

[27] G. D'Alfonso, *Intelligenza Artificiale e responsabilità civile. Prospettive europee*, in «Revista de Estudios Juridicos y Criminologicos», VI (2022), pp. 163-195.

tive measures to eliminate, or at least mitigate, the risks of damage caused by smart technologies, with a view to a multi-level liability mechanism[28].

Authors in favour of AI regulation, inspired by the principle of accountability, often start from the example provided by the regulation of personal data protection under the GDPR[29], due to the centrality of data in the functioning of AI[30]. However, when regulating civil liability for AI on the model of the GDPR[31], it would require the application of some precautions, in relation to the marked speciality that characterizes the regulation of personal data in the Euro-unitary area, which in turn also reverberates on the relationships between prevention rules and compensation rules[32]. In this regard, it is appropriate to keep the two dimensions, and consequently the functions, of responsibility separate in the context of AI regulation, while emphasizing the necessary dialogue that must exist between safety rules and liability rules to outline a more coherent framework of protections[33].

3. AI civil liability as an autonomous compensation case

The framing of the issue in the context of both product safety and data protection is justified by the approach adopted by the European AI regulatory institutions through legislation that is fully part of a new paradigm, which includes the protection of consumer health and safety at the heart of the system[34]; it is also consistent by virtue of the close relationship between the activity of processing personal data and AI, which led the European Commission to develop an AI Act that has significant features of continuity with the GDPR, especially with regard to the principle of accountability[35], which in the

[28] C. Faralli, *Diritti e nuove tecnologie*, in «Tigor», II (2019), pp. 43-52.

[29] EU Regulation 2016/679.

[30] *Ex multis*, G. Amore, *Fairness, Transparency e Accountability nella protezione dei dati personali,* in «Studium Iuris», IV (2020), pp. 414-429; G.M. Riccio, G. Scorza, E. Belisario, *GDPR e normativa privacy, Wolters Kluwer,* Milano 2018 pp. 236-250; M. Renna, *Sicurezza e gestione del rischio nel trattamento dei dati personali,* in «Responsabilità civile e previdenza», 85, IV (2020), pp. 1343-1363; G. Finocchiaro, *Intelligenza artificiale. Quali regole?,* Il Mulino, Bologna 2024, *passim.*

[31] F. Caroccia, *Ancora su responsabilità civile e uso delle intelligenze artificiali,* in «Contratto e impresa», II (2022), p. 424.

[32] T. De Mari Casareto Dal Verme, *cit.*, p.324.

[33] M.W. Monterossi, *Algorithmic Decisions and Transparency: Designing Remedies in View of the Principle of Accountability,* in «The Italian Law Journal», V, II (2019), pp. 711-730.

[34] S. Saccomani, *New frontiers of privacy, and consent in the information society*, in «EJPLT», II (2024), pp. 1-16.

[35] A. D'Alessio, *La responsabilità civile dell'intelligenza artificiale antropocentrica*, in «Persona e mercato», II (2022), pp. 243-257.

AI Act framework translates into the need to distribute responsibilities exactly along the value chain of smart products with a view to accountability, Cybersecurity and prevention[36].

The regulation of product safety follows a peculiar system logic in which the preventive function of accountability and the compensatory function of liability are kept distinct, but placed in correlation, to promote a protection framework that is as complete and multi-level as possible. From this point of view, the observation that entrusting the regulation of AI exclusively to civil liability would not provide suitable incentives for prevention is certainly correct. The rules of ex post liability alone are not sufficient to adequately incentivize prevention, especially in the field of the safety of smart products[37], for which the threat of compensation for damages does not exert a deterrent force comparable to the threat of criminal and administrative sanctions provided for at national level by the Consumer Code in the event of violation of safety regulations. Only in the first case, in fact, could the manufacturer accept the risk of compensating for the damage caused by the products he puts on the market and consequently decide to insure himself against any claims for compensation, transforming the risk of compensation into a business cost[38].

The difference in functions performed by accountability and liability is evident, so that the division of responsibilities is carried out more adequately through the rules of prevention, while the compensatory function typical of civil liability is of little interest in the internal distribution of costs among the various potential culprits, as at the centre of the system there is the interest of the injured party and the need to compensate him for the entire damage immediately regardless of the attribution to each defendant of the respective portion of liability. This approach emerges in the approach of the Product Liability Directive (PLD)[39], which explicitly provides for joint liability between the parties potentially involved in production, each of whom may act in recourse against the other co-responsible parties pursuant to Article 121 of the Consumer Code because of the size of the risk attributable to each[40].

[36] F. Serini, *La frammentazione del cyberspazio merceologico tra certificazioni e standard di cybersicurezza. Alcune considerazioni alla luce delle discipline europea e italiana*, in «Rivista informatica e diritto», I (2024), pp. 41-76.

[37] F. Busoni, *Sicurezza dei prodotti,* a cura di G. Vettori, *Codice del consumo. Commentario,* CEDAM, Padova, 2007, pp. 792-794.

[38] F. Ciccariello, *Sicurezza e qualità dei prodotti*, a cura di G. Recinto, M.L. Mezzaso, S. Certi, *Diritti e tutele dei consumatori,* Napoli, Edizioni Scientifiche Italiane, 2014, pp. 423-445.

[39] European Commision, *Proposal for a Directive of The European Parliament and of the Council on liability for defective products,* COM (2022)495 final, 28.9.2022.

[40] A.K. Mayrhofer, *Product liability in the age of AI-Proposal for a "two track" solution*, in «Revista electrónica de dereito», XXXIII, I (2024), pp. 106-127.

The damage caused by a defective product represents a single harmful event causally attributable to a potential plurality of jointly and severally liable, therefore, the causal attribution and the measurement of each contribution to the verification of the harmful event pertains to the phase of recourse between co-debtors, in which the risk of the unknown co-cause is made to weigh on the entire production chain[41].

Similar rules apply to smart and AI products, whose technological complexity requires the protection of injured parties by allowing the possibility of a unitary action that examines the elements of the case regardless of the exact identification of the specific responsibilities attributable to the components of the AI value chain[42]. The identification of individual economic responsibilities is a burden on the producer of the "composite product", since the consumer will be able to act against the final producer without having to bear the risk of uncertainty about one or more of the links in the production chain.

Where the final product, even with a high technological content, is the output of the production and assembly activity attributable to different parties, the legislation provides for joint liability[43]. The consumer suffers the damage regardless of the ascertainment of the supplier to whom the specific problem is attributable, and this is relevant only for the purposes of the internal division of responsibility. In this sense, the provision of ex ante safety rules on producers of AI systems and their components can be a criterion for quantifying the extent of regression between the different co-controllers[44].

The AI Act is characterized by a security by design approach, burdening each operator of AI systems with obligations defined, which take the form of an accountability of the algorithm, in which the technology itself can represent a useful substrate for dividing responsibilities among the parties involved by incorporating the rules on liability in the technologies, and vice versa[45].

The placement of goods and services on the market is capable of identifying types of considerable accurately predicted at the time of design, so that the product liability system does not find its distinctive character in the typification of legally relevant injuries by safety legislation; on the other hand, it does not appear economically efficient to encourage

[41] T. De Mari Casareto Dal Verme, *cit*., p.357.

[42] U. Carnevali, *Prodotto composto difettoso e regresso tra produttori responsabili. Il criterio delle «dimensioni del rischio»,* in «Responsabilità civile e previdenza», LXXX, II (2015), pp. 360-369.

[43] D. Cerini, V. Gorla, *Il danno da prodotto. Regole, responsabilità, assicurazione*, Giappichelli, Torino 2011, pp. 47-50.

[44] P. Bortone, L. Buffoni, *La responsabilità per prodotto difettoso e la garanzia di conformità nel codice del consumo*, Giappichelli, Torino 2007, pp. 80-83.

[45] A. Pajno, F. Donati, A. Perrucci, *Intelligenza artificiale e di ritto: una rivoluzione? Volume 2. Amministrazione, responsabilità, giurisdizione,* Il Mulino, Bologna 2022, p.173.

producers to adopt only the precautions provided for by the legislation ex ante, but they must also take into account the highest state of the art in order to place on the market safe products that meet consumers' safety expectations[46].

However, at a practical-application level, the reversal of the burden of proof, as a mechanism that should logically follow from an imputation criterion inspired by accountability, in the context of damage caused by products may not be effective since the injured party would only be responsible for proving the causal link between the AI system and the damage, while the defendant could free himself by demonstrating its compliance with safety legislation, de facto reducing consumer protection and inducing manufacturers to adopt a lower level of safety than traditional products[47].

The desire to treat AI safety and liability in the same way as the approach adopted so far by the European legislator in the field of product safety emerges precisely from the Commission's proposals, in which the preventive dimension of the accountability of the AI Act is kept distinct from the compensatory dimension of the proposals on civil liability, with the intention of creating coordination between different rules[48].

4. The European legal framework: the AI Liability Directive and the new Product Liability Directive

EU action on civil liability for AI has been divided into two different approaches[49]. One advocates the development of an ad hoc civil liability regime for AI; the other tends to harmonise certain circumscribed profiles. In the AI Act, providers and users of high-risk AI systems are primarily responsible ex ante, setting legal obligations related to the security of such systems and entrusting low-risk systems to a voluntary compliance basis If one wanted to establish liability for presumed fault with a reversal of the burden of proof, it

[46] E. Bellisario, *Il pacchetto europeo sulla responsabilità per danni da prodotti e da intelligenza artificiale. Prime riflessioni sulle Proposte della Commissione,* in «Danno e responsabilità», II (2023), pp. 153-166.

[47] T. De Mari Casareto Dal Verme, *cit.*, p.373.

[48] D. Amram, *Responsabilità, danno e accountability nella società dell'informazione,* in «Danno e responsabilità», G. Amore, *Fairness, Transparency e Accountability nella protezione dei dati personali*, in «Studium Iuris», IV (2020), pp. 414-429. (2019), pp. 243-247.

[49] E. Rajneri, *The projects of European regulation on liability for damage caused by artificial intelligence. Striking a balance between the interests of consumers, multinationals and SMEs*, in «Rivista di diritto bancario», II (2022), pp. 95-119.

would become of fundamental importance to identify the preventive obligations incumbent on operators, which would be largely set by the AI Act[50]. The purpose of the AI Liability Directive (AILD) [51], as defined by Article 1, concerns the simplification of certain evidentiary aspects in the legal systems of the Member States in cases of liability actions brought before national courts and based on the imputation of fault criterion, without affecting the contents of the relevant substantive law[52].

Such harmonisation would cover the disclosure of material information relating to high-risk AI systems, to enable the claimant in a liability judgment to adequately substantiate the claim for damages, and the burden of proof in tort civil actions brought before national courts to obtain compensation for damage caused by an AI system[53].

A system emerges inspired to reduce the information asymmetry existing between the injured party and the provider or user of an AI system, granting the former the right to obtain from the judicial authority an order for disclosure of information useful for initiating a possible compensation action or to access the burden of proof required by national rules on the subject in the case of a trial already instituted[54].

This provision, combined with the possibility of overcoming the presumption of causality by proving that the injured party could have access to useful information, would encourage suppliers and users to adopt the technical measures on registration imposed by the AI Act, even in the presence of critical issues in the field of industrial law. However, the main function of disclosure lies in ensuring substantial access to justice for the injured party, as this would allow the claimant to identify the responsible parties[55].

The Commission intends to carry out close coordination between the ex-ante security legislation, as set out in the AI Act, and the ex-post safety le-

[50] A. Bertolini, *Artificial Intelligence Does Not Exist! Defying the Technology Neutrality Narrative in the Regulation of Civil Liability for Advanced Technologies,* in «Europa e diritto privato», II (2022), pp. 369-420; G. Amore, *Fairness, Transparency e Accountability nella protezione dei dati personali,* in «Studium Iuris», IV (2020), pp. 414-429.

[51] Proposal for a Directive of the European Parliament and of the Council on adapting non-contractual civil liability rules to Artificial Intelligence (COM(2022) 496 fnal).

[52] U. Salanitro, *Intelligenza artificiale e responsabilità: la strategia della Commissione europea*, in «Rivista di diritto civile», VI (2020), pp. 1246-1276.

[53] N. Cappellazzo, *La AI liability directive: alcuni spunti di riflessione,* in «Federalismi.it», 28 dicembre 2022, pp. 1-5.

[54] C. Novelli, F. Casolari, A. Rotolo, M. Taddeo, L. Floridi, *Taking AI risks seriously: a new assesment model for the AI Act,* in «AI & Society», 39, V (2024), pp. 2493-2497.

[55] G. Noto La Diega, L.C.T. Bezerra, *Can there be responsible AI without AI liability? Incentivizing generative AI safety through ex-post tort liability under the EU*, in «International Journal of Law and Information Technology», XXXII (2024), pp. 1-21.

gislation, on civil liability, to hold accountable those who have specific obligations under the aforementioned legislation[56].

The AILD aims to complement the system of protections provided by the AI Act. Although the proposal specifies that it does not create substantive liability rules for AI, the current structure of the AILD seems to model a form of liability in which the imputation criterion is clearly skewed towards the non-compliance of the AI system with the AI Act, thus risking curbing the operation of national law on liability for fault[57].

The lack of incisiveness of the proposal is evident as it positivizes the judgment on causality and evidential simplification, requiring the injured party to prove the causal link between the output and the damage, and that the event that occurred constitutes the materialization of the risk that the violated precautionary rule was intended to prevent. These critical remarks could lead to a "flight" from AILD, in favour of rules that are already consolidated and often more flexible and protective than the mechanisms triggered by disclosure. As technological complexity increases, it becomes increasingly complicated to provide proof of human guilt, in which the criteria of imputation responding to different rationes take over[58].

The proposal for a new Product Liability Directive ("new PLD") has a twofold objective: to modernize the discipline contained in Directive 85/374/EEC, adapting some provisions to recent technological developments, and to provide evidentiary simplification mechanisms for those injured by technologically complex products, like what is envisaged in the AILD proposal. The first substantial difference with AILD lies in the broader scope of application of the PLD reform, which affects the changes brought about in the production reality by today's digital revolution[59]. The first substantial difference with the AILD lies in the broader scope of application of the PLD reform, which affects the changes brought about by today's digital revolution in the production environment.

The updated definition of "product" reflects the current complexity of the manufacturing world, including software and data for digital fabrication, whether marketed independently or as components of another product.

The Commission introduces the instrument of disclosure in Article 8, which can only be proposed during the proceedings. As regards the actual

[56] I (2024), pp. 113-121.

[57] S. Li, B. Sschütte, *The Proposed EU Artificial Intelligence Liability Directive: Does/Will Its Content Reflect Its Ambition?,* in «Technology and Regulation», I (2024), pp. 143-151.

[58] A. Mantelero, *The Fundamental Rights Impact Assessment (FRIA) in the AI Act: Roots, legal obligations and key elements for a model template,* in «Computer Law & Security Review: The International Journal of Technology Law and Practice», LIV (2024), pp. 1-18.

[59] B. Garcia, *The regulation of AI liability in Europe: a critical overview of two recent directive proposals–the (new) AILD and the (revised) PLD,* in «e-Publica», XI ,III (2024), pp. 28-41.

usefulness of disclosure, there is an analogy with the technical complexity of the data disclosed. In particular, in the context of the new LDP, the difficulty in identifying the information that needs to be released emerges. The directive does not specify who must demonstrate what the actual relevant information is, with the consequent danger that the producer will send the consumer even additional information, and conversely to create further litigation if the judges, in doubt, limit themselves to reproducing in the disclosure order the mere regulatory formula provided for in the directive[60].

The weakness of the presumptive mechanism emerges again, so that the producer, to avoid the presumption of defectiveness and take advantage of the ordinary evidentiary regime, will only need to disclose a plethora of information, leaving the plaintiff to manage a quantity of technically very complex data to discharge the ordinary burden of proof provided for by Article 9. The risk is that the plaintiff will be called upon to use the data obtained to base his claim for compensation and, therefore, to prove the defectiveness of the product. In essence, this right of access to relevant information would risk raising the standard to the use of complex technical data to support evidence that could instead be provided by means of simple presumptions based on the functioning of the product and the dynamics of the accident, thus frustrating the intention to simplify the burden of proof on the injured party[61].

However, such a provision risks compromising an important factor in the balance between the procedural positions of the parties, in which it is now a well-established principle in the Italian legal system that the risk of the unknown cause of the defect lies with the producer, while only the risk of the unknown cause of the damage falls on the consumer[62]. In so doing, the Commission appears to require the injured party to prove the technical component which gave rise to the defect precisely in those cases where it would be more difficult to provide it, placing the risk of the unknown cause of the defect on the consumer[63]. Although that increase in the burden of proof may be justified by the advantage which the injured party would receive from the neutralisation of one of the exculpatory evidence available to the producer,

[60] M. Lillà Montagnani, M.C. Najjar, A. Davola, *The EU Regulatory approach(es) to AI liability, and its Application to the financial services market*, in «Computer Law & Security Review: The International Journal of Technology Law and Practice», LIII (2024), pp. 1-19.

[61] T. De Graaf, *Liability for artificial intelligence in EU Law,* a cura di V. Mak, E. Tjong Tjin Tai, A. Berlee, Research Handbook in Data Science and Law, Cheltenham, Edward Elgar Publishing, 2024, pp. 117-136.

[62] A. Racki Marinkovic, *Liability for AI-related IP infringements in the European Union,* in «Journal of Intellectual Property Law and Practice», XIX, X (2024), pp. 784-792.

[63] M.C. Buiten, *Product liability for defective AI*, in «European Journal of Law and Economics», LVII, I (2024), pp. 239-273.

it is nevertheless necessary to prevent that exception from becoming a general rule in the demonstration of the defect.

5. Conclusions

The complexity of the advent of new technological phenomena requires the identification of elements that act as a substratum of the possible offences, to verify whether the current legislation can provide adequate answers to the various requests for protection. Civil liability legislation makes it possible to respond appropriately to the problem of liability for damage caused by AI[64]. It is necessary to rethink the traditional civil law categories, not so much by questioning their validity and replacing them with new ones, but to overcome the conception that these categories are immovable and rigid entities. The examination of the discipline and doctrine analysed has returned a hybrid and flexible imputation criterion that reflects the progressive overcoming of the traditional dichotomy between liability for fault and strict liability, enhancing the judicial role capable of shaping the nature of the rule applied to the concrete case, combining the essential cases with that rule on an operational level. This flexibility appears adequate and functional to the dynamism of current technological progress[65].

In particular, the possible risks related to the coordination carried out by AILD with the AI Act in terms of under-protection of injured parties were pointed out, to which the European legislator will have to pay attention so that the declared intention to provide a higher standard of protection for those damaged by AI systems than for those damaged by traditional products does not result in the creation of a liability subsystem that, In practice, it would raise the standard of proof and, therefore, complicate the plaintiff's procedural position[66]. Finally, the proposal for a new PLD fully reflects the supranational regulator's approach to the relationship between AI and civil liability. On the one hand, the proposal shows the need for legislative intervention also in the field of civil liability, regarding the updating of certain central notions in the construction of the directive with respect to the most recent advances in technology, which could represent a significant turning point in the regulation of

[64] G. Finocchiaro, *La regolazione dell'intelligenza artificiale,* in «Rivista trimestrale di diritto pubblico», LXXII (2022), pp. 1085-1099.

[65] I. Gredka-Ligarska, *In Search of Adequate Principles of AI Civil Liability*, in «PRAWO i WIEZ», III (2024), pp. 157-190.

[66] L. Heimerdinger, P. Patti, *Designing a European Liability Regime for AI. Questions and Challenges in the Aftermath of the AI Act*, in «Osservatorio del diritto civile e commerciale», XIII, II (2024), pp. 505-530.

AI. The criterion for attributing liability would remain unchanged, and the defect in its relational dimension would continue to represent the prerequisite for attributing liability to the producer[67]. Hence the relevance of national law which, albeit with the help of future legislative innovations, will see the interpretative activity of judges once again at the centre of the evolution of the matter. In conclusion, the national interpreter is called upon to interpret the evolutionary capabilities of smart products in the light of the principles that permanently govern the matter of tort, and which find their justification in the need to compensate for damages, reconciling the implementation of the right compromise between the protection of the injured and technological progress[68].

[67] T. De Mari Casareto Dal Verme, *cit.* p.392.

[68] N. TH. Nikolinakos, *Adapting the EU Civil Liability Regime to the Digital Age: Artificial Intelligence, Robotics, and Other Emerging Technologies.* Vol. 68. Springer Nature, Cham 2024, pp. 21-29.

BIBLIOGRAPHY

Amore G., *Fairness, Transparency e Accountability nella protezione dei dati personali*, in «Studium Iuris», IV (2020), pp. 414-429.

Amram D., *Responsabilità, danno e accountability nella società dell'informazione*, in «Danno e responsabilità», II (2019), pp. 243-247.

Astone A., *Sistemi intelligenti e regole di responsabilità*, in «Persona e mercato», 2, III (2023), pp. 593-595.

Bellisario E., *Il pacchetto europeo sulla responsabilità per danni da prodotti e da intelligenza artificiale. Prime riflessioni sulle Proposte della Commissione*, in «Danno e responsabilità», II (2023), pp. 153-166.

Bertolini A., *Artificial Intelligence Does Not Exist! Defying the Technology Neutrality Narrative in the Regulation of Civil Liability for Advanced Technologies*, in «Europa e diritto privato», II (2022), pp.369-420.

Bortone P., Buffoni L., *La responsabilità per prodotto difettoso e la garanzia di conformità nel codice del consumo*, Giappichelli, Torino 2007.

Botero Arcila B., *AI liability in Europe: How does it complement risk regulation and deal with the problem of human oversight?*, in «Computer Law & Security Review: The International Journal of Technology Law and Practice», LIV (2024), pp. 1-17.

Bovens M., *Two Concepts of Accountability: Accountability as a Virtue and as a Mechanism*, in «West European Politics», XXXIII V (2010), pp. 946-967.

Buiten M.C., *Product liability for defective AI*, in «European Journal of Law and Economics», LVII, I (2024), pp. 239-273.

Bussani M., *Le funzioni delle funzioni della responsabilità civile*, in «Rivista di diritto civile», II (2022), pp. 264-306.

Busoni F., *Sicurezza dei prodotti*, a cura di G. Vettori, *Codice del consumo. Commentario*, Padova, CEDAM, 2007, pp. 792-794.

Capilli G., *Responsabilità civile e Robot*, in «La nuova giurisprudenza civile commentata», III (2019), pp. 621-631.

Cappellazzo N., *La AI liability directive: alcuni spunti di riflessione*, in «Federalismi.it», 28 dicembre 2022, pp. 1-5.

Carnevali U., *Prodotto composto difettoso e regresso tra produttori responsabili. Il criterio delle «dimensioni del rischio»*, in «Responsabilità civile e previdenza», LXXX II (2015), pp. 360-369.

Caroccia F., *Ancora su responsabilità civile e uso delle intelligenze artificiali*, in «Contratto e impresa», II (2022), pp. 408-428.

Cassatella A., *La responsabilità funzionale nell'amministrare. Termini e questioni*, in «Diritto amministrativo», III (2018), pp. 677-746.

Callier M, Callier H., *Blame it on the machine: socio-legal analysis of liability in an ai world*, in «Washington Journal of Law, Technology & Arts»,XIV,I (2018), pp. 49-76.

Cerini D., Gorla V., *Il danno da prodotto. Regole, responsabilità, assicurazione*, Giappichelli, Torino 2011.

Chiti E., *L'accountability delle reti di autorità amministrative dell'Unione europea*, in «Rivi-

sta italiana di diritto pubblico comunitario», XXII I (2012), pp. 29-84.

Ciccariello F., *Sicurezza e qualità dei prodotti*, a cura di Recinto G., Mezzaso M. L., Cherti S., *Diritti e tutele dei consumatori*, Napoli Edizioni Scientifiche Italiane, 2014, pp. 423-445.

Coppini L., *Robotica e intelligenza artificiale: questioni di responsabilità civile*, in «Politica del diritto», IV (2018), pp. 713-740.

Comandè G., *Responsabilità ed accountability nell'era dell'Intelligenza Artificiale*, a cura di F. Di Ciommo, O. Troiano, *Giurisprudenza e autorità indipendenti nell'epoca del diritto liquido. Studi in onore di Roberto Pardolesi*, La Tribuna, Piacenza, 2018, pp.1001-1013.

Comandè G., *Intelligenza artificiale e responsabilità tra* liability *e* accountability*. Il carattere trasformativo dell'IA e il problema della responsabilità*, in «Analisi Giuridica dell'Economia», I (2019), pp. 169-188.

Comandè G., *Multilayered (Accountable) Liability for Artificial Intelligence*, a cura di S. Lohsse - R. Schulze - D. Staudenmayer, *Liability for Artificial Intelligence and the Internet of Things*, Baden-Baden, Nomos, 2019, pp. 165-184.

Comandè G., *Lettera sulla responsabilità (civile) e l'autonomia (individuale)*, in «Danno e responsabilità», VI 2022, pp. 665-668.

D'Alessio A., *La responsabilità civile dell'intelligenza artificiale antropocentrica*, in «Persona e mercato», II (2022), pp. 243-257.

D'Alfonso G., *Intelligenza Artificiale e responsabilità civile. Prospettive europee*, in «Revista de Estudios Juridicos y Criminologicos», VI (2022), pp. 163-195.

De Graaf T., *Liability for artificial intelligence in EU Law*, a cura di V. Mak E. Tjong Tjin Tai – A. Berlee, *Research Handbook in Data Science and Law,* Cheltenham, Edward Elgar Publishing, 2024, pp. 117-136.

De Mari Casareto dal Verme T., *Intelligenza artificiale e responsabilità civile. Uno studio sui criteri di imputazione*, Edizioni Scientifiche Italiane, Napoli, 2024.

Di Gregorio V., *Intelligenza artificiale e responsabilità civile: quale paradigma per le nuove tecnologie?*, in «Danno e responsabilità»,I (2022), pp. 51-63.

Di Rosa G., *Quali regole per i sistemi automatizzati intelligenti*, in «Rivista di diritto civile», V (2021), pp. 823-853.

Etzioni A., Etzioni O., *Keeping ai legal*, in «Vanderbilt Journal of Entertainment & Technology Law»,XIX ,I (2016), pp. 133-146.

Faini F., *Il diritto nella tecnica: tecnologie emergenti e nuove forme di regolazione*, in «Federalismi.it», XVI (2020), pp. 79-117.

Faralli C., *Diritti e nuove tecnologie*, in «Tigor», II (2019), pp. 43-52.

Finocchiaro G., *Intelligenza artificiale e protezione dei dati personali*, in «Giurisprudenza Italiana», VII (2019), pp. 1670-1677.

Finocchiaro G., *La regolazione dell'intelligenza artificiale*, in «Rivista trimestrale di diritto pubblico», LXXII,2022, pp.1085-1099.

Finocchiaro G., *Intelligenza artificiale. Quali regole?,* Il Mulino, Bologna 2024.

Gaeta M.C., *Liability rules and self-driving cars. The evolution of tort law in the light of new technologies,* Edizioni Scientifiche Italiane, Napoli 2019.

Garcia B., *The regulation of AI liability in Europe: a critical overview of two recent directive proposals—the (new) AILD and the (revised) PLD*, in «e-Publica», XI, III (2024), pp. 28-41.

Gatt L., Caggiano I., Gaeta M.C., *Italian Tort Law and Self-Driving Cars: State of art and Open Issues*, a cura di B.H. Oppermann, J. Stender-Vorwacs, *Autonomes Fahren. Technische Grundlagen, Rechtsprobleme, Rechstsfolgen*, München, Beck, 2020, pp. 239-255.

Gredka-Ligarska I., *In Search of Adequate Principles of AI Civil Liability*, in «PRAWO i WIEZ», III (2024), pp. 157-190.

Heimerdinger L., Patti F.P., *Designing a European Liability Regime for AI. Questions and Challenges in the Aftermath of the AI Act*, «Osservatorio del diritto civile e commerciale», XIII, II (2024), pp. 505-530.

Li S., Schütte B., *The Proposed EU Artificial Intelligence Liability Directive: Does/Will Its Content Reflect Its Ambition?*, in «Technology and Regulation»,I (2024), pp. 143-151.

Lillà Montagnani M., M.C. Najjar, A. Davola, *The EU Regulatory approach(es) to AI liability, and its Application to the financial services market*, in «Computer Law & Security Review: The International Journal of Technology Law and Practice», LIII (2024), pp. 1-19.

Mantelero A., *The Fundamental Rights Impact Assessment (FRIA) in the AI Act: Roots, legal obligations and key elements for a model*

template, in «Computer Law & Security Review: The International Journal of Technology Law and Practice», LIV (2024), pp. 1-18.

Marcello R., *Accountability: una prospettiva politologica*, in «Enti non profit», VII (2007), pp. 457-461.

Mayrhofer A.K., *Product liability in the age of AI — Proposal for a "two track" solution*, in «Revista electrónica de dereito», 33, I (2024), pp. 106-127.

Mazzamuto S., *Trattato di Diritto Privato*, Giappichelli, Torino 2021.

Monterossi M.W., *Algorithmic Decisions and Transparency: Designing Remedies in View of the Principle of Accountability*, in «The Italian Law Journal», V,II(2019), pp. 711-730.

Mulgan R., *Holding Power to Account: Accountability in Modern Democracies*, Palgrave Macmillan, London 2003.

Nikolinakos N. Th., *Adapting the EU Civil Liability Regime to the Digital Age: Artificial Intelligence, Robotics, and Other Emerging Technologies*. Vol. 68. Springer Nature, Cham 2024.

Noto La Diega G., Bezerra L.C.T., *Can there be responsible AI without AI liability? Incentivizing generative AI safety through ex-post tort liability under the EU*, in «International Journal of Law and Information Technology», XXXII (2024), pp. 1-21.

Novelli C., Casolari F., Rotolo A., Taddeo M., Floridi L., *Taking AI risks seriously: a new assesment model for the AI Act*, in «AI & Society», XXXIX, V (2024), pp. 2493-2497.

Pajno A., Donati F., Perrucci A., *Intelligenza artificiale e diritto: una rivoluzione? Volume 2. Amministrazione, responsabilità, giurisdizione*, Il Mulino, Bologna 2022.

Perlingieri P., Giova S., Prisco I. (a cura di), *Rapporti civilistici e intelligenze artificiali: attività e responsabilità*. Atti del 15° Convegno Nazionale 14-15-16 maggio 2020, Edizioni Scientifiche Italiane, Napoli, 2020.

Polverari L. *Alcune riflessioni sul l'accountability della politica di coesione europea, sul trade-off tra diversi tipi di accountability e sul trade-off tra accountability ed efficacia*, in «Rivista giuridica del Mezzogiorno», IV (2011), pp.1093-1108.

Racki Marinkovic A., *Liability for AI-related IP infringements in the European Union,* in «Journal of Intellectual Property Law and Practice», XIX, X (2024), pp. 784-792.

Rajneri E., *The projects of European regulation on liability for damage caused by artificial intelligence. Striking a balance between the interests of consumers, multinationals and SMEs*, in «Rivista di diritto bancario», II (2022), pp. 95-119.

Renna M., *Sicurezza e gestione del rischio nel trattamento dei dati personali*, in «Responsabilità civile e previdenza», LXXXV, IV (2020), pp.1343-1363.

Riccio G.M., Scorza G., Belisario E., *GDPR e normativa privacy*, Wolters Kluwer, Milano 2018.

Saccomani S., *New frontiers of privacy, and consent in the information society*, in «EJPLT», II (2024), pp. 1-16.

Salanitro U., *Intelligenza artificiale e responsabilità: la strategia della Commissione europea*, in «Rivista di diritto civile», VI (2020), pp. 1246-1276.

Salvi C., *Quali funzioni per la responsabilità civile?*, in «Danno e responsabilità», II (2022), pp. 137-140.

Serini F., *La frammentazione del cyberspazio merceologico tra certificazioni e standard di cybersicurezza. Alcune considerazioni alla luce delle discipline europea e italiana*, in «Rivista informatica e diritto», I (2024), pp. 41-76.

Teubner G., *Soggetti giuridici digitali? Sullo status privatistico degli agenti software autonomi*, E.S.I., Napoli 2019.

Thelisson E., Verma H., *Conformity assessment under the EU AI act general approach*, in «AI and Ethics», IV,I (2024), pp. 113-121

A. Vedder; L. Naudts, *Accountability for the use of algorithms in a big data environment*, in «International Review of Law, Computers & Technology», XXXI, II, (2017), pp. 206-224.

Vladeck D.C., *Machines Without Principals: Liability Rules and Artificial Intelligence*, in «Washinghton Law Review», LXXXIX (2014), pp. 117-149.

EDOARDO SALVATORE*

Minors and age verification systems: the protection of the «mobile born» in the information society§

Abstract

The paper is aimed at analysing the role of the child in the digital environment, and specifically on the Internet, an aterritorial place par excellence, where anyone has the possibility of interfacing with an indefinite number of users and content. This is, without any doubt, an extraordinary opportunity to foster the development of a person's identity, but at the same time it can present dangers that are often underestimated especially when those involved are particularly vulnerable individuals such as minors.The article offers an analysis about the processing of minors' data, which is one of the main innovations contained in EU Regulation No. 2016/679 (GDPR) and to which the European legislator has dedicated in Article 8 a specific regulation on the age of consent. The focus is on the issue of age verification systems in the online dimension, since age verification must take place in compliance with the aforementioned Regulation as well as with certain principles, including that of data minimisation. Therefore, some examples of age verification methods adopted by the most important online platforms are illustrated in order to demonstrate the need for balance between the safety of the child and the protection of the child's fundamental rights and freedoms.

Il contributo è volto ad analizzare il ruolo del minore nell'ambiente digitale, e specificamente in Internet, luogo aterritoriale per eccellenza, in cui chiunque ha la possibilità di interfacciarsi con un numero indefinito

* PhD student in "Diritto e Management della sostenibilità" at Siena University.

§ This contribution develops the presentation at the conference 'Cybersecurity e prodotti: sfide e responsabilità nelle società interconnesse',29 November 2024, Italian Supreme Court.

di utenti e contenuti. Si tratta, senza dubbio alcuno, di una straordinaria opportunità per favorire lo sviluppo dell'identità della persona, ma che al contempo può presentare pericoli spesso sottovalutati in particolar modo quando ad essere coinvolti sono soggetti segnatamente vulnerabili come i minori. L'articolo offre un'analisi circa il trattamento dei dati dei minori che costituisce una delle principali novità contenuta nel Regolamento UE n. 2016/679 (GDPR) e a cui il legislatore europeo ha dedicato all'art. 8 una puntuale disciplina sull'età del consenso. L'attenzione si concentra sul tema dei sistemi di age verification nella dimensione online, dal momento che la verifica dell'età deve avvenire nel rispetto del sopra citato Regolamento nonché di alcuni principi, tra cui quello di minimizzazione dei dati. Vengono, quindi, illustrati alcuni esempi di metodi per la verifica dell'età adottati dalle più importanti piattaforme online, al fine di dimostrare l'esigenza di equilibrio tra sicurezza del minore e protezione dei diritti e delle libertà fondamentali dello stesso.

1. The child as a vulnerable subject in the network

It is now well known that in the digital age characterized by an unstoppable revolution in communication and personal interaction[1], minors (and not only) spend a considerable part of their free time using electronic *devices* (television, computers, *smartphones*...) generating «a series of problems, including legal ones, not all of which have been adequately addressed and resolved in legislation»[2].

This paradigm shift, further amplified in the wake of the Covid-19 pandemic (which highlighted the need to use IT tools for everyday activities), calls for careful reflection on the role and position of minors in the digital environment[3].

Particularly significant in the development of the personality of the so-called «*mobile born*»[4], an expression that has replaced the previous epithet «digital natives»[5] is, without any doubt, the role of the Internet (or telematic «network»)

[1] On the new modes of socialisation from which real 'new languages' have developed, see A.I. Cesarano, I nativi digitali tra rischi e opportunità. Eu Kids online research, YouCanPrint, 2018.

[2] In these terms G. Magno, *Elementi di diritto minorile. La tutela dell'infanzia e dell'adolescenza nel diritto interno e internazionale*, Giuffrè, Milano 2019, p. 469.

[3] On the subject see G. Alpa, *I contratti del minore. Appunti di diritto comparato,* in «I Contratti», V (2004), pp. 517 ff., and more recently R. Senigaglia, *Minore età e contratto. Contributo alla teoria della capacità*, Giappichelli, Torino 2021.

[4] Mobile born' are those children who learn to use electronic devices at the same time or even before they take their first steps.

[5] A term coined by M. Prensky, *Digital Natives, Digital Immigrants,* in «On the Horizon», MCB University Press, IX (2001), pp. 1 ff. to refer to young people born after 1985, i.e. all those who have li-

that is increasingly accessible by everyone with every electronic device and in every place[6].

Indeed, on the Internet there is the circulation of an impressive amount of available information that, on the positive side, can be profitably used for the education of minors, while on the negative side it could represent a risk for the serene and complete development of their identity. In this sense, one need only think of all those sites that refer to illicit content or child pornography, access to which is extremely easy for anyone regardless of age[7].

The minor's action on the *web* is not limited to the possibility of operating as an Internet user or «passive surfer», who therefore accesses a range of content and information available on the web but includes the ability to navigate as an «active» who performs telematic acts, also in the context of e-commerce, qualifying as a true «consumer»[8].

In any case, the vulnerability *status*[9] that characterises the figure of the minor relates to the (in)ability to understand the potential contents with which he/she interfaces, rather than to the use of the technological means, in relation to which he/she often tends to show particular dexterity[10]. And it is precisely this familiarity in the use of *devices* with an Internet connection that justifies the analysis of the child-technology relationship according to an inverted criterion: the more skilled and safer the child is in surfing the net, the more he or she needs a high level of legal protection.

ved in close contact with digital technologies since birth, as opposed to 'digital immigrants' who, having become acquainted with the Net at a later stage of their lives, are immigrants to the digital world.

[6] Particularly explanatory are the data collected by Unicef, which estimate that approximately 175,000 children and young people access the Internet every day, with an average of one every half second; in global terms, one third of users are minors. Unicef, *On children's rights in relation to the digital environment: AGIA, CIDU and UNICEF Italy launch the new UN General Comment, 7* February 2022, in unicef.it.

[7] Thus E. Battelli, *Diritto privato delle persone minori di età*, Giappichelli, Torino, 2021, pp. 112 ff.

[8] In this regard F. Panuccio Dattola, *Minori e Internet*, Giappichelli, Torino 2009, pp. 48 ff., who defines the minor as the «object of an unconscious virtual contract». In contractual relations, the weak position of the minor consumer represents a real status; in other words, it would constitute a legally relevant and non-discontinuous condition capable of distinguishing the subject from others by reason of strictly individual peculiarities or for belonging to a specific group. For a more complete reconstruction of the notion of «consumer» see G. Alpa, Diritto dei consumatori in G. Alpa, A. Catricalà il Mulino, Bologna 2016, pp. 17 ff; while for the concept of «status» see P. Rescigno, voice *Status*, in «Enciclopedia giuridica I, Teoria generale», XXX, Treccani, Roma 1993, pp. 1 ff., who defines status as the practical instrument capable of creating or maintaining inequalities and 'zones of singular right' for subjects in a natural situation of infirmity such as minors.

[9] For a more in-depth examination C. Perlingieri, *La tutela dei minori di età nei "social network"*, in «Rassegna di diritto civile», IV (2016), pp. 1324 ff.

[10] See E. Andreola, *Misure cautelari a tutela dei minori nei social network,* in «Famiglia e diritto», IX (2021), pp. 850 ff.

The peculiarity of the Internet, but more specifically of the *social* community, is that it allows anyone to interact with an indeterminate number of users, often strangers, and with these to exchange messages, pictures, and, more generally, data[11]. The distorted sense of «intimacy»[12], the result of the erroneous conviction of sharing personal content exclusively with one's followers, puts minors in the position of making particularly sensitive aspects of their private lives public, contextually and irreversibly compromising their personal identity *in fieri*[13].

While it is evident, therefore, the vast range of potentialities deriving from such digital literacy[14] that can be exploited to build or entertain long-distance relationships with considerable opportunities for personality development, the risks and dangers that minors might incur cannot be underestimated, especially in view of their inexperience and fragility typical of their age[15].

The social networks that come into existence in the virtual space are not perceived by the child as something that is distinct from what develops in the physical world; on the contrary, the *web,* not being anything other than a different place where one can get to know and interact, becomes *in toto* a tangible extension of real life, indistinguishable from *onlife*[16].

Moreover, whenever one is dealing with a minor subject, the right question to ask is whether or not he or she can exercise the right or whether it is necessary to identify who exercises it in his or her name and on his or her behalf[17]. The Italian legislature has always considered the minor to be a subject deserving of particular protection, establishing in this sense a series of rules aimed at limiting the capacity to act, which according to the provisions of Article 2 of the Civil Code is acquired on reaching the age of majority[18].

[11] E. Battelli, *Minori e social network: cyberbullismo e limiti della parental responsibility,* in «Corriere giuridico», X (2021), pp. 1270 ff.

[12] On the subject A. Thiene, *I diritti della personalità dei minori nello spazio virtuale,* in «Annali online della didattica e della formazione docente», IX, XIII (2017), pp. 28 ff.

[13] On the topic T. Pasquino, *Identità digitale della persona, diritto all'immagine e reputazione, Privacy Digitale. Riservatezza e protezione dei dati tra GDPR e nuovo Codice Privacy,* a cura di E. Tosi, Milano, Giuffrè, 2019, pp. 93 ff. and G. Pino, *L'identità personale*, in *Trattato di Biodiritto*, a cura di S. Rodotà, P. Zatti, Giuffrè, Milano 2010, pp. 297 ff.

[14] On the topic F. Di Tano, *Minori, consenso "privacy" e vulnerabilità online: riflessioni alla luce del Regolamento generale sulla protezione dei dati (UE) 2016/679*, in «Notizie di Politeia», CXXXVI, (2019), p. 46.

[15] Cf. G. Magno, *cit.*, p. 471.

[16] In these terms L. Floridi, *The Onlife Manifesto*, Springer, Oxford, 2015, pp. 21 ff.

[17] On the point G. Capilli, *Minori in rete tra consenso e verifica dell'età. Analisi comparata e proposte di adeguamento al GDPR*, in «MediaLaws», I (2024), pp. 5 ff.

[18] *Ibid.*

However, the determination of such a limit overlooks all those hypotheses in which adolescents are capable of self-determination and, more generally, are able to understand the consequences of their actions, even though they have not yet reached the age of majority. Consequently, if the establishment of a single age threshold for the attribution of capacity to act runs the risk of excessively limiting the autonomy of minors in the real world, this effect is particularly amplified in the *online* dimension[19], where, all the more reason why it would be necessary to approach a graduation of the capacity of discernment in line with what is established by the national laws of the various Member States on the processing of minors' personal data[20].

The European legislator has intervened, in fact, by determining a precise regulation of the consent of the child relating «information society services»[21]. However, the European discipline emphasises the 'circulatory' rather than the «representative» moment of the person, without affecting it as it could have done[22].

2. The protection of minors in the European landscape: reflections in the light of the GDPR

The provision to be referred to is Article 8 of the GDPR[23], which is characterised by a limited scope of application that is already apparent from the heading; in fact, the direct reference to «information society services»[24] cir-

[19] In this regard M. Giandoriggio, *I minori d'età e i social network: l'insostenibile leggerezza del post,* in «Danno e responsabilità», 3, 2024, pp. 298 ff. who calls for an adaptation of the institution of the capacity to act, specifically declined to digital and the preparation of a statute that takes into account the specific capacity of discernment of the minor and the relative self-responsibility.

[20] G. Capilli, *Minori in rete tra consenso e verifica dell'età. Analisi comparata e proposte di adeguamento al GDPR, cit.*, p. 6.

[21] On the subject E. Battelli, *Il trattamento dei dati nel prisma della tutela della persona minore di età*, in «Diritto dell'informazione e dell'informatica», II (2022), pp. 270 ff.

[22] In this sense A. Thiene, *Segretezza e riappropriazione di informazioni di carattere personale: riserbo e oblio nel nuovo regolamento europeo,* in «Nuove leggi civili commentate», II (2017), p. 410.

[23] In fact, recital 38 of the Regulation already provides that '[children] deserve specific protection with regard to their personal data, as they may be less aware of the risks, consequences and interesting safeguards and of their rights in relation to the processing of personal data. This specific protection should, in particular, cover the use of children's personal data for marketing purposes or the creation of personality or user profiles and the collection of personal data relating to children when using services provided directly to a child'. The protection also extends to the manner in which communications are made to the child, in fact, recital 58 specifies that 'given that children deserve specific protection, when data processing concerns them, any information and communication should use plain and clear language that a child can easily understand'.

[24] For the definition of which please refer to Article 1(1)(b) of Directive (EU) 2015/1535.

cumscribes the scope of the latter, the basis of this intervention to modernise *the privacy* rules being the need to adapt the legislation on the subject to the current social context, in which minors make ever greater use of the various telematic services[25]. A form of autonomy in the management of themselves and their digital identity is therefore recognised for minors[26].

Paragraph 1 of Art. 8 GDPR enshrines the validity of consent[27] given personally by the child over the age of 16.

Otherwise, if the child is under sixteen, the processing of his or her personal data is only lawful if and to the extent that consent is given or authorised by the holder of parental responsibility[28].

In other words, the Community legislator has prepared a specific regulatory design in which the acquisition of capacity to act is anticipated, configuring a sort of 'digital age of majority'[29] and within which national legislators are granted discretion to identify a different age, albeit not lower than thirteen[30].

[25] In these terms F. Ruggeri, *Minors and Privacy, in Diritto privato delle persone minori di età,* a cura di E. Battelli, cit., p. 147.

[26] See F. Resta, *Conditions applicable to children's consent in relation to information society services,* in GDPR e normativa privacy. Commentario a cura di G.M. Riccio, G. Scorza, E. Bellisario, Ipsoa, Milano 2018, pp. 84-85.

[27] Consent is generally defined by the regulation as an 'unambiguous positive act by which the data subject manifests his or her free, specific, informed and unambiguous intention to accept the processing of personal data concerning him or her'. For a more complete examination of the issue, see I.A. Caggiano, *Il consenso al trattamento dei dati personali tra Nuovo Regolamento Europeo (GDPR) e analisi comportamentale. Initial insights*, in «Law Market Technology», (2017), pp. 7 ff.

[28] The structure of the rule makes it possible to identify the position and the degree of autonomy of the minor who benefits from network services on the basis of reaching a certain year of age: a distinction is therefore made between the so-called *petit enfants*, i.e. children in the strict sense for whom the logic of protection applies, and the so-called *grands enfants,* i.e. 'big minors' who instead enjoy greater freedom and self-determination. On this point, see F. Naddeo, *Il consenso al trattamento dei dati personali del minore*, in «Diritto dell'informazione e dell'informatica», I (2018), p. 27.

[29] On the difference between the age required to validly conclude a contract and the age required under the GDPR to consent to data processing, see P. Gallo, *Il consenso al trattamento dei dati personali come prestazione,* in «Rivista di diritto civile», VI (2022), p. 1054, who warns that «*under these conditions, a first possibility might be to consider that consent to the processing of personal data and that to the conclusion of the contract run on two different tracks, with the consequence that while consent to the processing of data could be given from the age of fourteen, consent to the conclusion of the contract of exchange could only be given after the age of eighteen; which could, moreover, give rise to application results that are, to say the least, absurd and impractical; indeed, the solution appears at first sight to be baroque and unnecessarily complex*».

[30] G. Scorza, *Appena nati e già online,* in *Specchio*, 8 January 2023, highlights how 'the European data protection regulation, the famous GDPR expressly states that under the age of sixteen – fourteen in Italy – a child is not able to give the provider of a digital service consent to process his or her data. But, of course, if that child claims to be older and the provider believes him, the problem is solved or, at least, online, it is often pretended to be so'.

This «digital maturity», to be qualified as an evolutionary achievement of 'non-adults' on the wave of a certain «juvenile protagonism»[31], on a par with the codified maturity referred to in Article 2 of the Civil Code, is subject to the automatism of a specific date, i.e. the fourteenth year of age, regardless of any assessment of the fundamental «capacity of discernment»[32] which is not only not mentioned, but which, on the contrary, has been identified as a guiding criterion in the areas of existential autonomy provided for by the Civil Code itself (e.g. Article 315-bis of the Civil Code) and which alone would justify such an anticipation.

On this point, it is also necessary to point out that, unlike the transition to majority, which presupposes the attainment of a suitable degree of maturity to be able to look after one's own interests, it is not very clear to which type of digital maturity the aforementioned digital maturity refers[33].

In any case, the choice adopted by the European Union legislature testifies to its renunciation of uniformity of discipline, given that the Member States, by using the different age options between thirteen and sixteen, have caused an inevitable fragmentation that does not benefit the process of European integration and not even minors, who are treated differently according to their nationality[34].

Moreover, the provision of an age threshold valid for all Member States would not have affected the legal rules on the validity, formation or effectiveness of contracts concluded by minors themselves[35] (Art. 8(3) GDPR).

On the basis of the above, it is clear that although today's legal framework offers multilevel protection from international to European sources

[31] In these terms B. Agostinelli, *Minori in rete: l'illusione del consenso e l'equivoco della "privacy"*, in «Familia», III (2023), pp. 405 ff.

[32] Cf. L. Bozzi, I *dati del minore tra protezione e circolazione per una lettura non retorica del fenomeno*, in «Europa e diritto privato», I (2020), p. 265 ff., who assumes the discernment of the 14-year-old child at the basis of his consent.

[33] B. Agostinelli, *Minori in rete: l'illusione del consenso e l'equivoco della "privacy", cit.*, p. 406.

[34] On the subject G. Capilli, La tutela dei dati personali dei minori, in Circolazione e protezione dei dati personali, tra libertà e regole di mercato. Commentary on the EU Regulation no. 2016/679 (GDPR) and the revised Legislative Decree no. 196/2003 (Privacy Code), a cura di R. PANETTA, Giuffrè, Milano 2019, pp. 258-259.

[35] Indeed, it is only in those hypotheses in which the processing of personal data constitutes, at least in part, the object of the service that there may be a need for coordination of the rules on the requirements of privacy consent with those relating to the prerequisites for the validity and effectiveness of the contract. On this point, see E. Battelli, *Il contratto di accesso ad Internet*, in «MediaLaws», I (2021), pp. 147 ff., who proposes the example of access to a telematic service that postulates the provision of consent to the processing of personal data for promotional purposes; in such a case, if the minor has reached the age of digital consent, but not the age required by national law in contractual matters, the lawfulness of the processing will be overturned by the exercise of the action for annulment of the contract.

of law[36], the effectiveness of the legal instruments available to children surfing the net is still not considered sufficient to guarantee an adequate level of protection[37].

In fact, the Regulation does not specify how in practice consent is to be given, nor how it is to be acquired[38].

Certainly, 'ticking' a box stating that one is at least sixteen years old is not a reasonable and effective method of providing adequate protection for minors, not to mention that this «flag» function can be performed by the most modern AI systems that are perfectly capable of circumventing such a superficial verification method[39].

Although there is no lack of different methods of age verification in the national and international panorama, such as, for example, that based on checking identity documents through images or videos used by gambling or virtual dating applications, a standardised system for verifying the legitimacy of consent given cannot yet be said to exist. If it is true that excessive bureaucratisation that could constitute an obstacle to the use of online services must be avoided, it is equally true that it is necessary to avoid falsification of one's age by minors, which would compromise the effectiveness of the Regulation[40].

3. Age verifications systems: a critical analysis of evolving concepts

Here, it seems useful to illustrate, albeit by way of example and with relative brevity, some of the methods adopted by the main platforms aimed at age verification.

One example is the ChatGPT case, in relation to which the Garante per la protezione dei dati personali (Garante for the protection of personal data) in its decision of 30 March 2023 noted the absence of an adequate tool for verifying the age of users, even though the service offered by OpenAI was reserved to people who had reached the age of thirteen[41]. Subsequently, the same

[36] See G. Pedrazzi, *Minori e social media: tutela dei dati personali, autoregolamentazione e privacy,* in «Informatica e diritto», I (2017), pp. 448 ff.

[37] See J. Adamovic', *Impact of General Data Protection Regulation on Children's Rights in Digital Environment,* in «Annals FLB-Belgrade Law Review», III LXIV (2016), p. 205.

[38] Very significant on this topic are the reflections of Working Party 29, *Guidelines on Consent under Regulation (EU) 2016/679,* 10 April 2018.

[39] On the subject G. Capilli, *Minori in rete tra consenso e verifica dell'età. Analisi comparata e proposta di adeguamento al GDPR, cit.*, pp. 8-9.

[40] *Ibid.*

[41] Specifically, according to the Garante, the processing of personal data of users of the service was contrary to Articles 5, 6, 8, 13 and 25 of the GDPR.

Guarantor ordered OpenAI to a) require all users connecting from Italy to pass an *age gate* in order to exclude underage users and b) submit to the Guarantor a plan for the adoption of age *verification* tools to exclude access to the service to all users under thirteen and underage users in the absence of a manifestation of will on the part of those exercising parental responsibility over them[42] . At present, a *box* has been included in the welcome screen reserved for Italian users who have already registered for the service, whereby, in order to access the service again, it is necessary to declare that they are over the age of majority or over thirteen and, in this case, have parental consent; in addition, the service registration mask has been set up to request the date of birth with a simultaneous block for users under thirteen and, in the case of users over thirteen (but minors), confirmation of parental consent to use the service[43].

A further noteworthy case is the one involving the famous social platform TikTok following the death of a ten-year-old girl from a *social challenge.* In that case, following its own investigation, the Garante had ordered the provisional limitation of the processing, with an express prohibition on the further processing of the data of users who connected from the Italian territory and for whom the personal data requirement was not certain. To this day, when registering for the social account, users are required to indicate their date of birth, a telephone number and the entry of a six-digit code previously sent by SMS when *signing up*[44]. It is therefore evident that there is a defect in the *ex ante* ascertainment by the platforms of the biographical requirement, which is noted on the basis of unilateral declarations by the users.

The third example that seems useful to recall has *Tencent*, one of the main Chinese technology companies operating in the field of video games, as its protagonist, and which, in order to limit gaming time, chose to use facial recognition to prevent children from going online at specific times[45]. In other words, there is enhanced monitoring for users who log on before eight o'clock and after twenty-two o'clock. Finally, *Tencent* requires the entry of one's national identification document, which is checked directly against the national database of citizens provided by the Ministry of Public Security. Given that mi-

[42] In these terms M. MARTONI, Systems of 'age verification'. Una prima esplorazione fra salvaguardia dell'interesse prioritario del minore e protezione dei dati personali, in «Ciberspazio e Diritto», III (2023), pp. 310 ff.

[43] *Ibid.*

[44] *Ibid.*

[45] This Chinese intervention is in the context of reducing so-called vamping, to be understood as a pathological phenomenon that generates insomnia due to addiction to electronic devices. For a more specific framing of the topic, see A.M. Della Vedova, L. Covolo, M. Muscatelli *et al.*, *Psychological Distress and Problematic Smartphone Use: Two Faces of the Same Coin? Findings from a Survey on Young Italian Adults,* in «Computers in Human Behavior», CXXXII, 2022.

nors could easily circumvent the control system, for instance, by using their parents' account, facial recognition has been set up, so that the system by means of a special algorithm will be able to analyse the data and *fully* comply with the principle of «maximum control»[46].

Added to this, however, is the critical issue of the additional data that are collected by *gaming* platforms and that would then be traced back to a specific subject and not also to a generic user[47].

The small number of cases illustrated demonstrates the need to ensure adequate protection of minors in relation to inappropriate content and at the same time the inefficiency of today's age verification systems based on mere unilateral declarations.

The *Tencent* case, on the other hand, demonstrated the possibility of actually verifying both the age and the identity of the person using a digital service, while at the same time making clear the «cost» – in terms of restricting rights and freedoms – that is deemed necessary[48].

The design of efficient *age verification* systems, however, must be conceived with a view to balancing the security and development of the data subject with his or her fundamental rights and freedoms, consistent with the core principles of minimisation and necessity.

The current legislation undoubtedly favours the minor's freedom of self-determination with respect to the right to protection of personal data, but at the same time leaves room for the realisation of harmful and irreversible consequences linked to his vulnerability. To all this must be added the not inconsiderable and worrying tendency of minors to relinquish control over their personal data in exchange for rewards[49] such as the use of services free of charge, in the name of the so-called *commodification* of personal data[50].

Generally speaking, in addition to the dubious effectiveness of consent as a tool legitimising the processing of personal data not only of minors, but the behavioural data also that testifies to a lack of awareness and rationality in personal privacy decisions becomes relevant[51].

[46] Thus M. Martoni, *Systems of "age verification". Una prima esplorazione fra salvaguardia dell'interesse prioritario del minore e protezione dei dati personali, cit.*, p. 313.

[47] This refers to data such as interactions, game modes or times, data collected for identity verification through facial recognition, or even system or technical data.

[48] *Ibid.*

[49] See I.A. Caggiano, *Il consenso al trattamento dei dati personali tra Nuovo Regolamento Europeo e analisi comportamentale, cit.,* p. 35.

[50] On the subject G. Pedrazzi, *Minori e social media: tutela dei dati personali, autoregolamentazione e privacy, cit.*, p. 451.

[51] On this subject, I.A. Caggiano, *Il consenso al trattamento dei dati personali tra Nuovo Regolamento Europeo e analisi comportamentale, cit.,* p. 35.

In conclusion, in the light of the critical issues outlined above and the lack of a ready and easily implemented technical-legal solution, is desirable to re-determine age thresholds and shared standards at European level that take into due consideration the discernment capacity and vulnerability of minors also and especially in the digital dimension[52].

As things stand, it is essential on the one hand to speed up the adoption of uniform political and legislative actions, and on the other hand to give a pivotal role to e-citizenship culture and education, since these are the only elements that can offer immediate protection and are independent of platform interests and regulatory interventions.

[52] Reference is made on this point to S. Livingstone, *Children: a special case for privacy?*, in «Intermedia», II (2018), p. 23.

BIBLIOGRAPHY

Adamovic' J., *Impact of General Data Protection Regulation on Children's Rights in Digital Environment*, in «Annals FLB-Belgrade Law Review», III, LXIV (2016), pp. 205-220.

Agostinelli B., *Minori in rete: l'illusione del consenso e l'equivoco della "privacy" in* «Familia», III (2023), pp. 397-415.

Alpa G., *Diritto dei consumatori*, a cura di G. Alpa, A. Catricalà, il Mulino, Bologna 2016.

Alpa G., *I contratti del minore. Appunti di diritto comparato*, in «*Contratti*» V (2004), pp. 517-526.

Andreola E., *Misure cautelari a tutela dei minori nei social network, in «Famiglia e diritto»,* IX (2021), pp. 848-868.

Battelli E., *Diritto privato delle persone minori di età*, Giappichelli, Torino 2021.

Battelli E., *Il contratto di accesso ad Internet, in «MediaLaws»*, I (2021), pp. 129-157.

Battelli E., *Il trattamento dei dati nel prisma della tutela della persona minore di età*, in «*Diritto dell'informazione e dell'informatica*», II (2022), pp. 267-301.

Battelli E., *Minori e social network: cyberbullismo e limiti della parental responsibility*, in «*Corriere giu*ridico», X (2021), pp. 1269-1277.

Bozzi L., *I dati del minore tra protezione e circolazione per una lettura non retorica del fenomeno*, in «*Europa e diritto privato*», I (2020), pp. 251-275.

Caggiano I.A., *Il consenso al trattamento dei dati personali tra Nuovo Regolamento Europeo e analisi comportamentale,* in «*Technology Market Law*», (2017), pp. 7-50.

Capilli G., *La tutela dei dati personali dei minori*, in *Circolazione e protezione dei dati personali, tra libertà e regole di mercato. Commentary on the EU Regulation no. 2016/679 (GDPR) and the revised Legislative Decree no. 196/2003 (Privacy Code)*, a cura di R. Panetta, Giuffrè, Milano 2019.

Capilli G., *Minori in rete tra consenso e verifica dell'età. Analisi comparata e proposte di adeguamento al GDPR*, in «*MediaLaws*», I (2024), pp. 1-24.

Cesarano I.A., *Digital natives between risks and opportunities. Eu Kids online research*, *YouCanPrint*, 2018.

Della Vedova A.M., Covolo L., Muscatelli M. *et al., Psychological Distress and Problematic Smartphone Use: Two Faces of the Same Coin? Findings from a Survey on Young Italian Adults*, in «*Computers in Human Behavior*», 132, 2022.

Di Tano F., *Minori, consenso "privacy" e vulnerabilità online: riflessioni alla luce del Regolamento generale sulla protezione dei dati (UE) 2016/679*, in «*Politeia News*», 136, 2019, pp. 46-56.

Floridi L., *The Onlife Manifesto*, Springer, Oxford, 2015.

Gallo P., *Il consenso al trattamento dei dati personali come prestazione*, in «*Rivista di diritto civile*», VI (2022), pp. 1054-1092.

Giandoriggio M., *I minori d'età e i social network: l'insostenibile leggerezza del post*, in «*Danno e responsabilità*», III (2024), pp. 296-306.

Livingstone S., *Children: a special case for privacy?*, in «*Intermedia*», 46 II (2018), pp. 18-23.

Magno G., *Elementi di diritto minorile. La tutela dell'infanzia e dell'adolescenza nel diritto interno e internazionale*, Giuffrè, Milano 2019.

Martoni M., *Age verification systems. Una prima esplorazione fra salvaguardia dell'interesse prioritario del minore e protezione dei dati personali*, in «*Ciberspazio e Diritto*», III (2023), pp. 307-322.

Naddeo F., *Il consenso al trattamento dei dati personali del minore*, in «*Diritto dell'Informazione e dell'informatica*», I (2018), pp. 27-64.

Panuccio Dattola F., *Minori e Internet*, Giappichelli, Torino 2009.

Pasquino T., *Identità digitale della persona, diritto all'immagine e reputazione*, in *Privacy Digitale. Confidentiality and data protection between GDPR and the new Privacy Code*, a cura di E. Tosi, Giuffrè, Milano 2019.

Pedrazzi G., *Minori e social media: tutela dei dati personali, autoregolamentazione e privacy*, in «*Informatica e diritto*», I (2017), pp. 437-451.

Perlingieri C., *La tutela dei minori di età nei "social network"*, in «*Rassegna di diritto civile*», IV (2016), pp. 1324-1340.

Pino G., *L'identità personale*, in *Trattato di Biodiritto*, a cura di Rodotà S., Zatti P., Giuffrè, Milano 2010.

Prensky M., *Digital Natives, Digital Immigrants*, in «*On the Horizon*, MCB University Press», IX (2001).

Rescigno P., voce *Status*, in «*Enciclopedia giuridica Treccani, I, Teoria generale*», XXX, Roma, 1993.

Resta F., *Condizioni applicabili al consenso dei minori in relazione ai servizi della società dell'informazione*, in *GDPR e normativa privacy. Commentario*, a cura di Riccio G.M., Scorza G., Bellisario E., Ipsoa, Milano 2018.

Ruggeri F., *Minori e privacy*, in *Diritto privato delle persone minori di età*, a cura di Battelli E., Giappichelli, Torino, 2021.

Scorza G., *Appena nati e già online*, in *Specchio*, 8 January 2023.

Senigaglia R., *Minore età e contratto. Contributo alla teoria della capacità*, Giappichelli, Giappichelli, Torino 2021.

Thiene A., *I diritti della personalità dei minori nello spazio virtuale*, in «*Annali online della didattica e della formazione docente*», XIII (2017), pp. 26-39.

Thiene A., *Segretezza e riappropriazione di informazioni di carattere personale: riserbo e oblio nel nuovo regolamento europeo*, in «*Nuove leggi civili commentate*», II (2017), pp. 410-444.

VIVIANA VISAGGIO*

NIS 2: the European Strategy on Cybersecurity between notification obligations and impact on stakeholders§

Abstract

The European Cybersecurity strategy has undergone rapid evolution in recent years, driven by the growing importance of digital security in an increasingly interconnected world. The European Union has recognised the need to protect critical infrastructure, personal data and digital economies from growing cyber threats by adopting a series of regulations aimed at strengthening the security of information systems[1]. The first significant step in this direction was the NIS Directive, adopted in 2016, which required Member States to improve the security of networks and information systems in key sectors such as energy, transport, healthcare and banking. However, the evolution of threats has required an update, leading to the adoption of the NIS 2 Directive in 2022. This new directive has broadened the scope of the previous one, including new strategic sectors and strengthening mandatory security measures for businesses and public administrations.

The NIS 2 Directive introduces more stringent obligations for the notification of Cybersecurity incidents. So-called essential entities must follow a rigorous three-step notification process: a preliminary report within 24 hours of identifying the incident, an update within 72 hours and a final report within one month. This mandatory process is similar to that

* PhD student at University of Salamanca. Researcher at the Research Centre of European Private Law (ReCEPL).

§ This contribution develops the presentation at the conference 'Cybersecurity e prodotti: sfide e responsabilità nelle società interconnesse',29 November 2024, Italian Supreme Court.

1 F. Cerciello, *NIS2 Directive: how to prepare for the new cyber security objectives,* Cybersecurity360, 2023, available at https://www.cybersecurity360.it/cybersecurity-nazionale/direttiva-nis-2-ecco-come-prepararsi-a-recepire-inuovi-obiettivi-di-cyber-security

provided for in the GDPR for the notification of data breaches, and emphasises the importance of adequate organisation for the management of cyber incidents. Cooperation between Member States is also strengthened through the creation of a cooperation network for the coordinated management of incidents. In addition to financial penalties, the Directive introduces an additional element: organisations affected by significant cyber incidents will be required to publish information about the penalties on their own channels, exacerbating the damage to their reputation.

This transparency requirement, together with financial penalties, highlights how the consequences of a cyber incident are not limited to direct costs, but also extend to reputational losses and crisis management costs. To avoid these consequences, it is essential that those involved take appropriate preventive measures, improving Cybersecurity and strengthening data analysis and incident management procedures. The NIS2 Directive, together with the national transposition legislation – Law 90/2024[2] and the very recent Legislative Decree 138/2024[3] – reflect the European Union's desire to address cyber threats in a coordinated and systematic manner, ensuring the protection of critical infrastructure and the continuity of essential services.

La strategia europea in materia di Cybersicurezza ha conosciuto una rapida evoluzione negli ultimi anni, trainata dalla crescente rilevanza della sicurezza digitale in un mondo sempre più interconnesso. L'Unione europea ha riconosciuto la necessità di proteggere le infrastrutture critiche, i dati personali e le economie digitali dall'aumento delle minacce informatiche, adottando una serie di misure regolatorie volte a rafforzare la sicurezza dei sistemi informativi.

Il primo passo significativo in questa direzione è stato compiuto con l'adozione della direttiva NIS nel 2016, che ha imposto agli Stati membri di migliorare la sicurezza delle reti e dei sistemi informativi in settori chiave quali l'energia, i trasporti, la sanità e il settore bancario. Tuttavia, l'evoluzione delle minacce informatiche ha reso necessario un aggiornamento del quadro normativo, culminato nell'adozione della direttiva NIS 2 nel 2022. Tale direttiva ha ampliato l'ambito di applicazione della discipli-

[2] Law No. 90 of 28 June 2024 Provisions on strengthening national Cyber security and cybercrime. (24G00108) (Official Journal General Series No. 153 of 02-07-2024).

[3] Legislative Decree No. 138 of 4 September 2024 Transposition of Directive (EU) 2022/2555 on measures for a high common level of Cybersecurity across the Union, amending Regulation (EU) No. 910/2014 and Directive (EU) 2018/1972 and repealing Directive (EU) 2016/1148. (OJ General Series No. 230 of 01-10-2024).

na previgente, includendo nuovi settori strategici e rafforzando gli obblighi di sicurezza a carico delle imprese e delle pubbliche amministrazioni.

La direttiva NIS 2 introduce obblighi più stringenti in materia di notifica degli incidenti di Cybersicurezza. Le cosiddette entità essenziali sono tenute a rispettare un articolato procedimento di notifica in tre fasi: una segnalazione preliminare entro 24 ore dalla conoscenza dell'incidente, un aggiornamento entro 72 ore e una relazione finale entro un mese. Tale meccanismo obbligatorio richiama quello previsto dal GDPR per la notifica delle violazioni dei dati personali e sottolinea l'importanza di un'adeguata organizzazione interna per la gestione degli incidenti informatici. È altresì rafforzata la cooperazione tra gli Stati membri mediante l'istituzione di una rete di cooperazione per la gestione coordinata degli incidenti. Accanto alle sanzioni pecuniarie, la direttiva introduce un ulteriore elemento di deterrenza: le organizzazioni colpite da incidenti informatici significativi sono tenute a rendere pubbliche, attraverso i propri canali, le informazioni relative alle sanzioni irrogate, con un conseguente aggravamento del danno reputazionale.

Tale obbligo di trasparenza, unitamente alle sanzioni economiche, evidenzia come le conseguenze di un incidente informatico non si esauriscano nei costi diretti, ma si estendano anche alle perdite reputazionali e ai costi di gestione della crisi. Per evitare tali ricadute, è essenziale che i soggetti coinvolti adottino adeguate misure preventive, rafforzando i livelli di Cybersicurezza e migliorando le procedure di analisi dei dati e di gestione degli incidenti. La direttiva NIS 2, unitamente alla normativa nazionale di recepimento – la legge n. 90/2024 e il recentissimo decreto legislativo n. 138/2024 – esprime la volontà dell'Unione europea di affrontare le minacce informatiche in modo coordinato e sistematico, garantendo la protezione delle infrastrutture critiche e la continuità dei servizi essenziali.

1. The European Cybersecurity strategy

In recent years, the European Union's Cybersecurity strategy has evolved rapidly, driven by a growing awareness of the need to protect critical infrastructure, personal data and digital economies from increasingly sophisticated and pervasive threats. The EU has adopted a series of regulatory measures aimed at strengthening the resilience and security of information and network systems, which are essential to the functioning of modern society.

Digital transformation and the increasing interconnection between individuals, businesses and institutions, amplified by cross-border exchanges, have placed information systems at the centre of everyday life. However, this rapid digitalisation has exposed the European digital ecosystem to a proliferation of cyber threats, requiring coordinated, effective and innovative responses from Member States.

The increase in the number, scope and complexity of cyber attacks, combined with their growing frequency, poses a serious threat not only to the functioning of technological infrastructures, but also to economic and social stability in the European Union's internal market. Cybersecurity incidents can seriously compromise economic activities in the internal market, causing significant financial losses, undermining user confidence and generating significant economic and social damage to European citizens.

In this context, the adoption of European legislation, such as the NIS 2 Directive, and the strengthening of the powers of the Cybersecurity Agency at national level, represent decisive steps towards the construction of a robust and uniform European Cybersecurity framework capable of responding to current challenges and ensuring the protection of a constantly evolving digital market.

Therefore, the European Union feels more than ever that it is essential to address Cybersecurity and its effectiveness for the better functioning of the internal market. Furthermore, Cybersecurity is a key enabler for many critical sectors to successfully implement digital transformation and reap the full economic, social and sustainable benefits of digitalisation[4].

The regulatory framework for the European digital single market initially took shape with EU Regulation No. 2016/679, known as the GDPR, which governs the protection of personal data in companies, and EU Regulation No. 2019/881, the Cybersecurity Act, which focuses on certifications for technological products and services. These are complemented by EU Directive No. 2016/1148, the so-called NIS Directive, now repealed and replaced by Directive (EU) 2022/2555 of 14 December 2022 of the European Parliament and of the Council (NIS Directive 2), which dealt with the security of networks and information systems for Operators of Essential Services (OES) and Digital Service Providers (DSPs).

The update of this legislative framework became necessary following a review of the effectiveness of the NIS Directive, which revealed a context of uncertainty for Cybersecurity at European level. The European Commission, through two impact assessments of the NIS, empha-

[4] Recital 3. EU Directive 2022/2555 of the European Parliament and of the Council of 14 December 2022.

sised the urgency of regulatory adaptation[5]. The Covid-19 pandemic further highlighted the critical issues of the NIS, as hospitals, often excluded from its scope, did not benefit from specific security measures[6]. In fact, the introductory report to NIS 2 makes it clear from the outset that Cybersecurity has also been recognised as a key element in the Commission's response to the crisis caused by the COVID-19 pandemic. The package includes a new Cybersecurity strategy aimed at strengthening the Union's strategic autonomy, improving collective resilience and promoting a global and open internet.

The Commission's assessments revealed insufficient resilience in EU businesses, uneven awareness among national institutions and a lack of adequate measures to deal with crises. Only 27% of small businesses, 51% of medium-sized businesses and 72% of large businesses had a Cybersecurity policy in place, making a more coordinated approach to strengthening the resilience of critical infrastructure, as set out in the 2020 EU Cybersecurity Strategy, a matter of urgency.

At the end of 2022, a legislative package was introduced to clarify and harmonise Cybersecurity regulations. This includes EU Regulation No. 2022/2554 (DORA), which establishes digital operational resilience requirements for the financial sector, and EU Directive No. 2022/2557 (CER), which strengthens the resilience of critical entities, replacing the 2008 directive on critical infrastructure. The NIS 2 Directive coordinates with these instruments, aiming at integrated cyber risk management and requiring Member States to adopt national strategies to identify and protect essential entities in eleven strategic sectors[7]. In an ever-changing digital landscape, the European Union is demonstrating a concrete commitment to strengthening Cybersecurity as a fundamental pillar for the protection of citizens and businesses as well as essential services such as healthcare. By adopting an increasingly comprehensive and integrated regulatory framework, the EU aims not only to mitigate existing vulnerabilities, but also to promote a secure, resilient and digital ecosystem capable of seizing the opportunities offered by global technological transformation.

[5] G. Marino, *Cybersecurity nell'UE: ecco le ragioni di una NIS2,* Agenda Digitale, (2022) available at https://www.agendadigitale.eu/sicurezza/cybersecurity- comune-nell-ue-ecco-le-ragioni-di-una-nis2/

[6] Proposal for a Directive Of The European Parliament And Of The Council on measures for a high common level of Cybersecurity across the Union, repealing Directive (EU) 2016/1148.

[7] C. Ogriseg, *Network and Information Security: la strategia europea nella Direttiva NIS II e le sfide che ci attendono,* « Ciberspazio e diritto», vol. 25, n. 76 I (2024), pp. 133-150.

2. The NIS 2 Directive

The first major initiative in this area was the NIS Directive (Network and Information Systems Directive), known as NIS1, adopted in 2016, which required Member States to improve the security of networks and information systems in strategic sectors[8].

Directive (EU) 2016/1148 was based on Article 114 of the Treaty on the Functioning of the European Union (TFEU), with the aim of promoting the internal market through the harmonisation of national regulations. However, the Cybersecurity requirements imposed on companies providing services or carrying out economic activities vary considerably between Member States, both in terms of the type of obligations and the methods of control. These differences generate additional costs and obstacles for companies operating across borders. In some cases, the requirements imposed by one country may conflict with those of another, further complicating cross-border activities. Furthermore, the uneven application of Cybersecurity obligations in one country can adversely affect the level of security in other countries, especially given the frequency of international trade.

The review of the Directive revealed wide differences in its application, including with regard to its scope, which was largely left to the discretion of each country. The Directive gave Member States considerable leeway in implementing security and incident notification obligations, which were therefore applied differently at national level. Similar differences can be found in the way the Directive's provisions are monitored and enforced.

The differences between Member States created fragmentation in the internal market and hindered its smooth functioning, especially with regard to the cross-border provision of services and the level of cyber resilience, due to the application of different measures. These divergences increased the vulnerability of some Member States to cyber threats, with potential consequences for the entire European Union. NIS 2 modernises the existing legal framework, taking into account the increasing digitalisation of the internal market in recent years and the rapidly evolving landscape of Cybersecurity threats[9].

The NIS 2 (Network and Information Security) Directive also finds its legal basis in Article 114 of the Treaty on the Functioning of the European Union

[8] F. Cerciello, *Direttiva NIS2: ecco come prepararsi a recepire i nuovi obiettivi di cyber security*, Cybersecurity360, 2023, available at https://www.cybersecurity360. it/cybersecurity-nazionale/direttiva-nis-2-ecco-come-prepararsi-a-recepire-inuovi- obiettivi-di-cyber-security/

[9] Proposal for a Directive of the European Parliament and of the Council on measures for a high common level of Cybersecurity across the Union, repealing Directive (EU) 2016/1148 Brussels 16.12.2020 of the European Commission.

(TFEU), which gives the European legislator the power to harmonise national regulations in order to ensure the smooth functioning of the internal market. This legal basis is justified by the need to eliminate divergences between Member States' internal regulations, which may hinder the free movement of goods and services and create disparities in the regulation of competition. The directive aims to standardise measures relating to cyber risk management and incident notification obligations, reducing disparities that adversely affect the cross-border activities of businesses and the full exercise of fundamental freedoms, such as the freedom of establishment and the freedom to provide services within the European single market.

The regulatory intervention complies with the principles of subsidiarity and proportionality, as the transnational nature of cyber threats makes exclusively national or regional solutions inadequate, requiring a coordinated approach at Union level. The NIS 2 Directive aims to overcome the shortcomings of the previous regulatory framework, which emerged during the transposition and implementation of the original NIS Directive, by strengthening the coordination and effectiveness of the measures taken at Member State level. In this sense, the Union's intervention aims to ensure a uniform level of security of networks and information systems, while promoting strategic cooperation between Member States.

The provisions comply with the principle of proportionality, as they do not go beyond what is necessary to achieve the objectives set. Although the measures envisaged involve initial implementation costs, they are proportionate to the expected benefits, such as reducing the economic and social damage resulting from possible cyber incidents and improving the overall resilience of critical infrastructure.

The choice of regulatory instrument fell on a directive, which was considered appropriate for striking a balance between harmonising rules and allowing the flexibility needed to take account of national specificities. Through the directive, Member States will be able to adapt the regulatory framework to their own needs, identifying essential and important entities in addition to those already provided for, without compromising the common rationale of ensuring a high and uniform level of Cyber security within the Union[10].

With the repeal of the previous directive, the sectoral scope is broadened to cover a larger part of the economy, ensuring protection for sectors and services that are critical to the social and economic activities of the internal market. The new directive overcomes the limitations of the old distinction

[10] Proposal for a Directive of the European Parliament and of the Council on measures for a high common level of Cybersecurity across the Union, repealing Directive (EU) 2016/1148 – E.C.

between operators of essential services and digital service providers, which is now considered outdated as it no longer reflects the real importance of sectors and services for the economic and social activities of the Union[11].

Under the previous Directive (EU) 2016/1148, each Member State was responsible for identifying entities that met the criteria for being considered essential service operators. In order to reduce the large differences between Member States and ensure legal certainty regarding Cybersecurity risk management measures and reporting obligations, it is preferable to introduce a uniform criterion for determining which entities fall within the scope of the Directive . This criterion should be based on a size threshold, covering medium-sized enterprises (as defined in Article 2(1) of the Annex to Recommendation 2003/361/EC) or enterprises exceeding the limits for medium-sized enterprises in that paragraph, which operate in the sectors and provide the services or perform the activities specified in the Directive. Member States should also consider including certain small and micro-enterprises, as defined in paragraphs 2 and 3 of Article 2 of that Annex, where these play a crucial role for society, the economy or specific sectors or services[12].

NIS 2 brings a significant change to the scope, overcoming the previous distinction between 'operators of essential services' and 'digital service providers' provided for in NIS and introducing uniform criteria for defining public and private operators, now divided into two new categories of critical entities: 'essential entities' and 'important entities'. This revision broadens the scope of NIS2 to include not only private operators in sectors considered critical by the European Union (such as energy, transport, finance, healthcare, digital infrastructure and drinking water), but also digital service providers (such as e-commerce platforms, search engines, cloud computing, social media and ICT management), public administration and space sector operators.

The directive also extends its scope to 'other critical sectors', regulating providers of publicly available communication networks and electronic communication services, domain name registration service providers, certain public administration bodies and other operators identified as critical. EU Member States will have until 17 April 2025 to draw up a list of essential and important entities that will provide the required information, with at least biennial updates and reviews, in order to ensure greater uniformity in the application of NIS2.

A size threshold is also introduced, whereby only medium-sized and large enterprises fall within the scope of NIS2, as defined in Article 2(1) of Recommen-

[11] A high common level of Cybersecurity – NIS2 – European Parliament report – available at https://www.europarl.europa.eu/RegData/etudes/ATAG/2022/738184/EPR.

[12] A high common level of Cybersecurity NIS 2' available at https://www.europarl.europa.eu/RegData/etudes/ATAG/2022/738184/EPR.

dation 2003/361/EC. Micro and small enterprises are excluded, with certain exceptions, such as where they are sole providers of electronic communications networks or publicly available electronic communications services, trust service providers, top-level domain name registries or public administration bodies.

The innovation mentioned above is particularly significant, as NIS2 will apply not only to organisations falling under the new definition of 'critical entities' of medium and large size, but also to the chain of subcontractors and service providers that support them, causing a domino effect[13].

A further extension is the inclusion of a territorial dimension: the new rules will apply to all entities providing services to critical sectors within an EU Member State, regardless of the legal seat of those entities (principle of extraterritoriality). This means that even companies based outside the EU could be subject to the NIS2 Directive if they provide services to critical entities in a Member State, even if they do not have offices within the European Union. This approach, similar to the strategy adopted by the GDPR for the protection of personal data, aims to broaden the scope of the provisions and ensure a common level of security within the EU.

NIS 2 aims to reduce existing disparities by establishing minimum standards for a coordinated regulatory framework, promoting cooperation between the competent authorities of each Member State, updating the sectors and activities subject to Cybersecurity obligations, and introducing redress mechanisms and enforcement measures to ensure the effective application of these obligations. Therefore, Directive (EU) 2016/1148 is repealed and replaced by the new directive[14].

The NIS 2 Directive identifies eleven strategic sectors considered essential for Cybersecurity and informat s in the European Union. These sectors are listed in Annex I to Directive (EU) 2022/2555 of the European Parliament and of the Council of 14 December 2022. The eleven strategic sectors defined by the NIS 2 Directive for Cybersecurity in the European Union are: Energy, Transport, Banking, Financial Market Infrastructure, Healthcare, Drinking Water, Wastewater, Digital Infrastructure, Public Administration, Space, Food. These sectors are considered crucial to EU society and the economy and must adopt security measures to protect networks and information systems, thereby contributing to the overall resilience of European critical infrastructure[15]. Furthermore, Legislative

[13] EPRS I European Parliament Research Service – Author: Mar Negreiro, Research Service for Members PE 738.184 – November 2022.

[14] https://digital-strategy.ec.europa.eu/ European Digital Strategy

[15] Directive (EU) 2022/2555 of the European Parliament and of the Council of 14 December 2022 on measures for a high common level of Cybersecurity across the Union, amending Regulation

Decree No. 138 of 4 September 2024, which transposes the NIS 2 Directive into Italian law, confirms these strategic sectors[16].

3. Entities subject to NIS 2 obligations: sectors and parameters

The NIS 2 Directive, transposed in Italy by Legislative Decree No. 138 of 2024[17] significantly broadens the scope of application compared to the previous legislation, including a wider range of sectors and entities. This expansion aims to ensure uniform protection against cyber threats, promoting cooperation between the competent authorities of Member States and introducing renewed implementation measures to ensure the effective application of the obligations provided for.

In particular, the new directive overcomes the distinction between 'operators of essential services' and 'digital service providers' in the previous NIS Directive, introducing uniform criteria for identifying obligated entities.

To determine whether a company falls within the scope of the legislation, three fundamental criteria can be used: territorial, sectoral and size.

The first criterion is territorial. The Directive applies to all companies providing products and services within the European Union, but its scope is not limited to entities with their registered office in a Member State. Article 2 introduces a principle of extraterritoriality, stipulating that the obligations established by NIS 2 also apply to companies located outside the EU, provided that they offer services or sell products on the European market. This mechanism ensures that global companies operating in critical sectors are subject to the Cybersecurity standards set out in the legislation, preventing non-European entities from evading their obligations simply because of their geographical location.

The second criterion is sectoral, identifying specific categories of economic activities considered critical to the security and stability of the Union. The Directive distinguishes between two groups of sectors listed in Annexes I and II. Annex I lists the eleven sectors considered to be highly critical, including energy, transport, banking, healthcare, public administration, drinking

(EU) No 910/2014 and Directive (EU) 2018/1972 and repealing Directive (EU) 2016/1148 (NIS Directive 2) available at https://eur-lex.europa.eu/legal-content/IT/TXT/?uri=CELEX%3A32022L2555

[16] The legislation on the National Cybersecurity Agency is available at https://www.acn.gov.it/portale/nis/la-normativa

[17] Legislative Decree No. 138 of 4 September 2024, implementing Directive (EU) 2022/2555 on measures for a high common level of Cybersecurity across the Union, amending Regulation (EU) No 910/2014 and Directive (EU) 2018/1972 and repealing Directive (EU) 2016/1148. (24G00155) available on the website (Official Journal General Series No. 230 of 01-10-2024).

water supply and wastewater management. Strategic sectors also include business-to-business ICT services, financial market infrastructures, the space sector and digital infrastructures. The latter category includes a wide range of operators, including cloud computing service providers, and is not limited to digital platforms alone, but extends to anyone who processes personal data and operates in the digital sector. Annex II further expands the scope of the Directive to include seven other critical sectors, including postal and courier services, waste management, chemical production and distribution, the food sector, transport equipment manufacturing, computer and medical device manufacturing, digital service providers and the research sector.

The third criterion is size-related and provides that the Directive applies to companies that exceed a certain size threshold. To fall within the scope of NIS 2, a company must have at least 50 employees and an annual turnover or balance sheet exceeding £10 million. However, the legislation makes a further distinction between obligated entities, identifying two categories: essential entities and important entities. Article 3 establishes that organisations exceeding the threshold for medium-sized organisations and operating in the sectors identified in Annex I, as well as large organisations and public administrations, are considered essential. Organisations operating in the sectors listed in Annexes I and II that do not fall within the category of essential organisations are defined as important organisations.

The extension of the subjective scope of application of NIS 2 compared to the previous NIS Directive is significant and reflects the European Union's growing focus on protecting critical infrastructure and essential services from increasingly sophisticated cyber threats. The new regulatory framework requires obligated entities to adopt rigorous security measures and strengthen their resilience to potential attacks, while also imposing stringent obligations in terms of risk management and incident reporting. In this context, the adoption of an impact measurement approach inspired by models already present in other EU regulations, such as the General Data Protection Regulation (GDPR), is particularly important. Article 35 of the GDPR introduces the obligation to carry out a data protection impact assessment (DPIA) for processing operations that are likely to result in a high risk to the rights and freedoms of data subjects. A similar principle is taken up by NIS 2 with the introduction of the so-called 'Failure Reporting and Impact Assessment' (FRIA), provided for in Article 9 of the Directive, which requires obligated entities to carry out an impact analysis on security incidents, assessing the risk with specific measurement tools.

The integration of impact assessment models such as the GDPR's DPIA and NIS 2's FRIA highlights a structured, risk-based approach that goes beyond mere regulatory compliance and requires a proactive assessment of the vulnerability of digital infrastructures and the measures needed to mitigate risks. Companies falling within the scope of the Directive must therefore equip themselves with appropriate risk measurement tools that enable them to identify potential threats, analyse attack scenarios and adopt timely response strategies.

In Italy, Legislative Decree No. 138/2024 defines the scope of application in Article 3, including the public and private entities listed in Annexes I, II, III and IV, which fall under national jurisdiction. Annexes I and II identify highly critical sectors (e.g. energy, transport, healthcare, digital infrastructure and drinking water) and other critical sectors (such as financial services, space infrastructure, food supply and waste management) respectively. Annex III lists the categories of public and private administrations concerned (including ministries, national research bodies and large healthcare institutions). Annex IV includes additional types of relevant entities, such as local public transport service providers, educational institutions engaged in research, cultural entities and affiliated or controlled companies, subject to specific identification by the competent authorities[18].

The NIS 2 Directive also introduces a territorial dimension, applying to all entities providing services to critical sectors within an EU Member State, regardless of their legal domicile. This approach, similar to that adopted by the GDPR, aims to broaden the scope of the provisions and ensure a common level of security within the European Union.

In conclusion, the NIS 2 Directive represents a significant step towards strengthening Cybersecurity in the European Union, extending obligations to a wider range of entities and sectors, and promoting greater cooperation between Member States to address growing cyber threats.

4. The supply chain and the importance of Cybersecurity: obligations and responsibilities for system resilience

The supply chain represents the set of steps necessary to bring a product or service from the initial supplier to the end customer. This process involves various actors, including manufacturers, raw material suppliers, transporters, warehouses, distributors and retailers. Each stage adds value or fa-

[18] https://www.acn.gov.it/portale/nis/la-normativa

cilitates the flow of goods and information, from the collection of raw materials to the processing, distribution and delivery of the final product[19].

The interconnection of numerous entities makes the supply chain vulnerable to cyber risks. To mitigate these risks, the European Union's NIS2 Directive imposes specific Cybersecurity obligations on strategic entities, including:

• Providers of essential and critical services, such as companies operating in the energy, transport, finance and healthcare sectors, where a disruption of services could have a significant impact on society and the economy[20].

• Digital service providers, including entities offering electronic communications services, cloud computing platforms, social media, e-commerce and other online services. Even small businesses in this sector, which are often involved in the management of sensitive data, must take appropriate security measures[21].

• Communication infrastructure and network operators, such as those who manage public networks and domain name registries, which are essential for the continuity and security of communications, crucial for the functioning of other critical services[22].

• Cyber security and ICT service providers, i.e. companies that provide technologies and services for the security management of other entities, such as firewalls, monitoring and incident response. These entities are an integral part of the overall security of the supply chain[23].

• Companies identified as critical, regardless of their size, which according to national assessments are of economic or strategic importance. These too must take measures to ensure adequate protection against cyber risks[24].

The adoption of such measures for the identified categories of entities aims to ensure that every link in the supply chain contributes to overall security, preventing vulnerabilities that could compromise cyber resilience at national and international level.

[19] J.M. Boyens *et al.*, Cybersecurity Supply Chain Risk Management for Systems and Organisations, 2022.

[20] https://www.acn.gov.it/portale/nis/la-normativa

[21] https://digital-strategy.ec.europa.eu/

[22] https://www.acn.gov.it/portale/nis/la-normativa

[23] Adaptation to the Decree implementing NIS 2 of 17 October 2024 available at https://www.osservatorio-dataprotection.it/tmt/digital-ict/ /

[24] National Strategic Framework for Cyberspace Security – Presidency of the Council of Ministers – December 2013.

5. Notification obligations in the event of cyber incidents

The new provisions contained in the NISc2 Directive significantly improve the European Union's ability to prevent, manage and respond to large-scale Cybersecurity incidents and crises. This is achieved through the introduction of clearly defined responsibilities, advanced planning and greater cooperation between Member States. The Directive requires each country to designate national authorities responsible for managing cyber crises, to develop large-scale national plans to deal with Cybersecurity incidents and to participate in the European Cyber Crisis Liaison Network (EU-CYCLONe). This network, established to support the coordinated management of incidents and crises at the operational level, is a key pillar of the EU framework outlined by the European Commission in its 2017 recommendation on coordinated response to Cybersecurity incidents and crises[25].

In Italy, the NIS2 Directive was transposed into national law by Law No. 90 of 2024, which updated the national regulatory framework on Cybersecurity. This process was completed by Legislative Decree No. 138/2024, which introduced enhanced measures for digital infrastructure and essential services. Among the main changes are specific responsibilities for businesses and public administrations in terms of risk management and obligations to report significant cyber incidents[26].

The NIS2 Directive requires critical entities to report cyber incidents that could cause serious operational disruption or financial loss. The reporting process follows a three-step procedure:

1. An initial notification within 24 hours of becoming aware of the incident.

2. An update within 72 hours providing further details on the incident.

3. A final report within one month, with a full description of the event and the measures taken.

This mandatory notification system is inspired, as mentioned above, by the mechanism provided for in the GDPR for data breaches[27]. The directive also emphasises the importance of having an adequate organisation in place to manage such incidents, providing for the strengthening of IT security and the implementation of specific procedures and roles to ensure a timely and effective response.

[25] Available at https://ec.europa.eu/

[26] Obligations available on the website https://acn.gov.it

[27] NIS2 Directive: new developments in the security of networks and information systems in the European Union, available at https://www.osservatorio-dataprotection.it/tmt/ direttiva-nis2-novita-in-tema-di-livello-di-sicurezza-delle-reti-e-dei-sistemi-informativi-nellunione-europea/

A European network called CyCLONe (Cyber Crisis Liaison Organisations)[28] is tasked with supporting the coordinated management of large-scale Cybersecurity incidents, as established by Article 16 of the NIS 2 Directive. This network facilitates the reporting of cyber incidents and collaborates with CSIRTs, the Cybersecurity incident response teams already established by the previous NIS Directive, which are responsible for receiving and handling reports. The new legislative framework in Italy expands the functions of the CSIRT and introduces provisions on cyber crisis management, with the definition of the relevant plan and the institutionalisation of the EU Cyber Crises Liaison Organisation Network (EU-CyCLONe) for the coordinated cross-border management of large-scale Cybersecurity crises at the operational level.

It is clear that a cyber incident can have negative repercussions on the integrity and security of personal data. Consequently, while recognising the autonomy of the GDPR with regard to the application of online privacy legislation, it is essential to strengthen organisational structures to manage such incidents effectively. In fact, among the recommended organisational measures, NIS 2 emphasises the need to appoint a Cybersecurity officer, clearly define the roles and responsibilities of personnel involved in incident management, and establish specific procedures to be followed in the event of a data breach[29].

6. Penalties and publication obligations

Compliance with the NIS 2 Directive is not limited to formal demonstration of regulatory compliance, but requires the actual and effective implementation of Cybersecurity measures. This approach reflects the European Union's intention to transform cyber risk management into an operational element integrated into corporate governance. However, this obligation could pose a si-

[28] The European *Union Cyber Crises Liaison Organisation Network* (EU-CyCLONe) is a cooperation network for national authorities in Member States responsible for managing cyber crises. The network was launched in 2020 and formalised following the entry into force of NIS2.

[29] Internal procedures for assessing incidents (events that compromise the availability, authenticity, integrity or confidentiality of stored, transmitted or processed data or of the services offered by or accessible through information and communication technology systems and networks) and near misses (events that could have compromised the availability, authenticity, integrity or confidentiality of stored, transmitted or processed data or of the services offered by or accessible through information and communication technology systems and networks) must be carefully evaluated.integrity or confidentiality of data stored, transmitted or processed or of services offered by or accessible through IT and network systems, but which was effectively prevented or did not occur) and cyber threats (any circumstance, event or action that could damage, disrupt or otherwise negatively impact the network and information systems, users of such systems and other persons). Thus, F. Bavetta, op. cit.

gnificant challenge, especially for small and medium-sized enterprises, which may find it difficult to bear the management costs given the complexity of the measures required[30].

A relevant aspect is the strengthening of the powers of supervisory authorities, which can intervene both preventively (ex ante) and after (ex post) cyber incidents. The penalties are severe: up to €10 million or 2% of annual turnover for 'critical entities' and up to €7 million or 1.4% of turnover for 'essential entities' ( ) (Article 34(5) of the NIS 2 Directive)[31]. These measures highlight the seriousness of the economic consequences for non-compliant organisations, but could also excessively penalise companies with limited economic resources, increasing the competitive gap between large enterprises and SMEs.

Another critical element concerns the responsibility of management bodies. The Directive takes an extremely procedural approach and requires not only the adoption of security measures, but also the maintenance of adequate awareness of cyber risks through continuous training of those identified as responsible and supervision of the practices adopted (Article 20 of the NIS 2 Directive entitled Governance)[32].This 'awareness' obligation is innovative and takes the form of a set of rules of conduct that nevertheless raise questions about the actual ability of the authorities responsible to verify the level of competence of directors and the fairness of individual sanctions, especially in complex business contexts.

Furthermore, the obligation for sanctioned organisations to make breaches public through their official channels introduces an additional dimension of reputational risk. Although such transparency may encourage companies to improve their defences, it also risks publicly exposing vulnerabilities that could be exploited by malicious competitors, for example. This requirement poses a dilemma between the need to inform the public and the need to preserve the operational security of the organisation[33].

Finally, the notification system introduced, inspired by the GDPR, provides for timely and detailed notifications in three stages (within 24 hours, 72 hours and one month). This model demonstrates a high level of attention to crisis management, but could be burdensome for less

[30] Special Report of the European Court of Auditors 21/2024: State aid in times of crisis – The response was swift, but there are shortcomings in the Commission's monitoring and inconsistencies in the framework supporting EU industrial policy objectives available at https://www.eca.europa.eu/it/publications/SR-2024-21

[31] Available at osservatorio-dataprotection.it

[32] National Cybersecurity Agency available at acn.gov.it

[33] Reputational risk and supply chain due diligence available at https://nhabi.it/rischio-reputazionale-e-due-diligence-sulla-supply-chain/

structured companies, especially if they are not supported by a rapid response infrastructure[34].

In conclusion, although the NIS 2 Directive represents a step forward for Cybersecurity at European level, it raises questions about the economic sustainability of the measures required, the balance between transparency and confidentiality, and the ability of SMEs to adapt without compromising their competitiveness. These aspects require careful national implementation and targeted support to prevent the obligations under NIS 2 from becoming an excessive burden for some entities, thereby undermining the objectives of inclusiveness and comprehensive protection.

7. A reflection on the implications for those involved

The growing incidence of cyber attacks has highlighted the need to update the European regulatory framework on Cybersecurity. The European Union has responded to this need through the NIS 2 Directive, which represents a significant evolution from the previous NIS Directive (2016/1148/EU), broadening its scope and introducing more stringent requirements for stakeholders. This regulatory update comes amid growing uncertainty about the Cybersecurity of European companies, with a particular focus on small and medium-sized enterprises operating in sectors that are strategic for Member States[35].

The NIS 2 Directive aims to provide greater clarity on the security measures to be adopted and to empower the management of obligated entities by introducing specific obligations for the adoption of technical and organisational measures proportionate to the risks. Voluntary certification of Cybersecurity systems is also strengthened as a tool to promote trust and digital resilience[36].

However, the NIS 2 Directive does not merely introduce new Cybersecurity obligations, but also has significant economic, reputational and operational implications for obligated entities. Compliance with these provisions is therefore not only a security requirement, but also a matter of economic sustainability and competitiveness for companies operating in strategic sectors.

[34] Personal data breaches (data breaches) according to the provisions of Regulation (EU) 2016/679 available at https://www.garanteprivacy.it/data-breach

[35] M Pellerzi, A. Michinelli, *Cyber Resilience Act, what changes for the security of digital products and IoT,* available at https://www.cybersecurity360.it/legal/cyber-resilience-act-cosa-cambia-per-la-sicurezza-dei-prodotti-digitali-e-iot-luci-e-ombre/

[36] National Cybersecurity Agency available at acn.gov.it

From an economic point of view, the penalties for non-compliance are particularly high. For essential entities, the Directive provides for fines of up to €10 million or 2% of global annual turnover, while for important entities, penalties can reach €7 million or 1.4% of global annual turnover[37]. These amounts represent a significant burden, particularly for medium-sized companies, which may find themselves having to allocate significant financial resources to compliance, taking them away from other areas of investment. This means that Cybersecurity is no longer just a matter of IT protection, but a critical variable in the financial management of companies, affecting their capacity for growth and innovation.

From a reputational perspective, the Directive imposes transparency obligations that can expose companies to consequences far more serious than financial penalties alone. Entities that suffer cyber incidents and significant security breaches are required to disclose the penalties they have received and notify the competent authorities in a timely manner[38]. The loss of trust on the part of customers, partners and suppliers can permanently compromise a company's competitiveness, especially in sectors where information security is a key factor in the relationship of trust with users, such as healthcare, banking and critical infrastructure. If made public, a cyberattack can have a negative impact on market positioning and affect the ability to attract new customers and investors, amplifying the damage suffered far beyond the initial event.

From an operational point of view, the implications of NIS 2 translate into the need to ensure the operational continuity and resilience of information systems, preventing an attack from compromising the provision of essential services to the community. Cybersecurity incidents can cause the interruption of essential services, such as transport, energy and healthcare, with indirect costs related to system restoration, crisis management and possible compensation to those affected[39]. In addition to the direct damage suffered by the entity affected by a cyberattack, a highly significant factor is the possible domino effect along the supply chain. Organisations that do not comply with security obligations can become the weak link in the entire network, compromising the stability of suppliers and upstream entities, causing slowdowns and disruptions on a much larger scale. Cybersecurity, therefore, no longer concerns only the individual involved, but must be considered from a systemic perspective, involving all interconnected actors within a given digital infrastructure.

To mitigate these consequences, it is essential to adopt a preventive approach based on risk measurement, as already provided for in the Failure

[37] Directive (EU) 2022/2555, Art. 34, paras. 4 and 5.
[38] Directive (EU) 2022/2555, Art. 23.
[39] Directive (EU) 2022/2555, Articles 21 et seq.

Reporting and Impact Assessment (FRIA) introduced by Article 9 of the Directive. This methodology, in line with the impact assessment required by the GDPR for the processing of personal data[40], requires the adoption of periodic risk assessment tools, the implementation of appropriate security measures, continuous staff training and the preparation of a detailed incident response plan. All this inevitably entails significant costs, which require consideration of who should bear the burden of this transition to greater digital security.

This consideration leads to a broader reflection on the relationship between Cybersecurity, economic power and access to technological resources. Cyberattacks can originate from anywhere on the global network, but they most easily target the weakest links in the chain: vulnerable individuals and less protected systems. In a context where compliance with security standards requires substantial investment, there is a risk that businesses with fewer economic resources will find themselves at a disadvantage compared to large infrastructures with advanced tools and greater financial capacity. Digital transformation must not, therefore, result in a mechanism of exclusion, in which security becomes a factor of discrimination between more structured entities and those that cannot afford the costs of compliance. The balance between technology, power and financial resources therefore becomes a central issue in the debate on Cybersecurity, so that protective measures do not become tools for further economic abuse, but are accessible and sustainable for all businesses, regardless of their size.

The debate on Cybersecurity needs to extend beyond the technical sphere to include a broader reflection on its social, ethical and political effects. Although defence against cyber threats is essential to ensure the continuity of critical infrastructure and essential services, it must be properly balanced with the needs of economic sustainability.

Cyber threats can come from anywhere on the global network and often target the weakest links in the chain, i.e. the most vulnerable individuals and the least protected IT systems[41]. The digital age requires continuous analysis of the relationship between technology and power, so that Cybersecurity tools do not themselves become a means of abuse by the 'richest' against companies with fewer economic resources.

[40] Regulation (EU) 2016/679, Art. 35.

[41] AGID Guidelines on the creation, management and storage of electronic documents May 2021 available at https://www.agid.gov.it/sites/default/files

BIBLIOGRAPHY

Bertolini A., *Cybersecurity Supply Chain Risk Management for Systems and Organisations*, NIST, 2022.

Cerciello F., *NIS2 Directive: how to prepare for the new cyber security objectives*, (2023) in «Cybersecurity360».

Marino G., *Cybersecurity nell'UE: ecco le ragioni di una NIS2*, (2022) in «Agenda Digitale».

Ogris Eg C., *Network and Information Security: la strategia europea nella Direttiva NIS II e le sfide che ci attendono*, in «Ciberspazio e diritto», I (2024), pp. 133-150.

Pellerzi M., Michinelli A., *Cyber Resilience Act: cosa cambia per la sicurezza dei prodotti digitali e IoT*, (2023) in «Cybersecurity360».

LAVINIA VIZZONI*

The "digital minors" between vulnerabilities and opportunities for risk containment§

Abstract

The rapid diffusion of digital technologies has profoundly reshaped the living environment of minors, generating unprecedented opportunities but also significant risks for their personal development and autonomy. Minors are increasingly exposed to Artificial Intelligence systems, interconnected devices, video games, and social platforms, often without adequate supervision and with potential vulnerabilities that may compromise their capacity for self-determination. While the EU Artificial Intelligence Act acknowledges the need to safeguard fundamental rights, it offers only limited and indirect protections specifically targeting minors, despite the heightened risk implications arising from their use of high-risk or minimally regulated AI systems. This gap becomes particularly evident in the domain of AI-enabled video games, where opaque data practices, dark patterns, profiling, behavioural manipulation, and unsafe interactions with other users expose minors to privacy violations, Cybersecurity threats, and possible grooming or cyberbullying. The interplay between the AI Act, the GDPR, the Digital Services Act, and sector-specific legislation illustrates a fragmented regulatory framework that struggles to effectively address the complex vulnerabilities of young users. In this context, certification schemes and Cybersecurity standards emerge as promising preventive tools capable of identifying, mitigating, and managing risks from the product design stage. An integrated approach to safety, combining product security, data protection, and Cybersecurity by design, appears essential to ensure the effective protection and well-being of digital minors within an increasingly interconnected technological ecosystem.

* Professoressa Associata di Diritto privato, Università degli Studi di Pisa.

§ This contribution develops the presentation at the conference 'Cybersecurity e prodotti: sfide e responsabilità nelle società interconnesse', 29 November 2024, Italian Supreme Court.

La rapida diffusione delle tecnologie digitali ha profondamente trasformato l'ambiente di vita dei minori, generando opportunità senza precedenti ma anche rischi significativi per il loro sviluppo personale e la loro autonomia. I minori sono sempre più esposti a sistemi di intelligenza artificiale, dispositivi interconnessi, videogiochi e piattaforme social, spesso senza un'adeguata supervisione e con potenziali vulnerabilità che possono compromettere la loro capacità di autodeterminazione. Sebbene la legge dell'UE sull'intelligenza artificiale riconosca la necessità di salvaguardare i diritti fondamentali, offre solo protezioni limitate e indirette specificamente rivolte ai minori, nonostante le maggiori implicazioni di rischio derivanti dal loro utilizzo di sistemi di IA ad alto rischio o minimamente regolamentati. Questa lacuna diventa particolarmente evidente nel campo dei videogiochi basati sull'IA, dove pratiche opache in materia di dati, modelli oscuri, profilazione, manipolazione comportamentale e interazioni non sicure con altri utenti espongono i minori a violazioni della privacy, minacce alla sicurezza informatica e possibili casi di adescamento o cyberbullismo. L'interazione tra la legge sull'IA, il GDPR, la legge sui servizi digitali e la legislazione specifica del settore illustra un quadro normativo frammentato che fatica ad affrontare efficacemente le complesse vulnerabilità dei giovani utenti. In questo contesto, i sistemi di certificazione e gli standard di sicurezza informatica emergono come strumenti preventivi promettenti in grado di identificare, mitigare e gestire i rischi fin dalla fase di progettazione del prodotto. Un approccio integrato alla sicurezza, che combini la sicurezza dei prodotti, la protezione dei dati e la sicurezza informatica fin dalla progettazione, appare essenziale per garantire la protezione efficace e il benessere dei minori digitali all'interno di un ecosistema tecnologico sempre più interconnesso.

1. "Digital Minors" Between Risks and Opportunities

A recent, very relevant factor, which has prompted a profound debate concerning the role and protection of the person under the age of 18, is the rapid development and consequent massive diffusion of scientific and technical knowledge and applications[1].

[1] On the relationship between law and science, see authoritatively OPPO G., Scienza, diritto, vita umana, in «Riv. dir. civ.»,I (2002), p. 11 who points out that applied science, and thus technology, is ontologically destined to be regulated by law. On this topic see also, equally authoritatively, ALPA G., Tecnologie e diritto privato, in «Riv. it. sc. giur. », (2017), p. 205 ff.

Indeed, a real new dimension has now opened up, giving rise to an equally new, further chapter in the discipline of juvenile and family law in general[2]: virtual reality enters the family and constitutes now an inseparable aspect of it[3].

Thus, a new specific profile emerges, which we could define as inherent to "digital minors": minors who commonly make use of applications that employ digital technologies, sometimes supervised and/or supported by their parents, sometimes instead in complete autonomy.

The minor in the digital context then becomes an even more problematic and articulated figure, because he or she represents, in a sense, the user par excellence of certain content: from accessing and surfing the Net, to accessing social networks, perhaps through smartphones, and again, to the daily use of devices whose operation is based on Artificial Intelligence mechanisms, moreover interconnected within the Internet of Things[4], such as wearables[5], smart home speakers[6] or smart toys[7], up to

[2] Talks about a new dimension" of family law, A.C. Di Landro, *Best interest of the child e tutela dei minori nel dialogo tra legislazione e giurisprudenza*, in «Nuove leggi civ. comm.», II (2020), p. 452 ff. On the relationship between new technologies and family law and the theorization of a kind of "cyberfamily," see exemplarily S. Nardis, *La famiglia e gli affetti nell'era digitale*, Naples, 2020, p. 7 ff. and with specific regard to children, p. 39 ff.

[3] See F. Di Ciommo, *Evoluzione tecnologica e regole di responsabilità civile,* Naples, 2003, 32 ff. On the massive diffusion of new technologies in the daily lives of minors as well, see also E. Andreola, *Minori e incapaci in Internet*, Naples, 2019, p. 22 ff.

[4] For a general overview of the notion of IoT and its legal implications see R.H. Weber, *Internet of Things-Need for a New Legal Environment?,* in «Computer Law & Security Review», (2009), p. 521 ff. On the potential of the IoT see A. Santosuosso, *Intelligenza artificiale e diritto,* Milan, 2020, p. 180 ff. On the applications of the IoT, see the concerns expressed by G. Ramaccioni, *La protezione dei dati personali e il danno non patrimoniale. Studio sulla tutela della persona nella prospettiva risarcitoria,* Naples, 2017, p. 16 ff. and p. 288 ff. See also F. Giovanella, *Le persone e le cose: la tutela dei dati personali nell'ambito dell'Internet of Things,* in V. Cuffaro, R. D'Orazio, V. Ricciuto, (eds.), *I dati personali nel diritto europeo*, Giappichelli, Torino, 2019, p. 1213 ff.

[5] See Italian Data Protection Authority Order No. 179 of March 26, 2015, Launching the Public Consultation on the Internet of Things, which states, «*The devices concerned are not only traditional computers or smartphones, but also those embedded in everyday objects ("things"), such as wearable devices (so-called wearables), home automation (domotics), and geo-referencing and assisted navigation; this entails the collection and management of data relating to the behavior, habits, preferences and health status of often unconscious users, with the effect of enabling their identification, either directly or indirectly, through the creation of even detailed profiles.*».

[6] On this point, may I refer to L. Vizzoni, *Domotica e diritto. La Smart Home tra regole e responsabilità*, Milan, 2021, p. 72 ff.

[7] The news has reported sensational cases, such as the affair of the Hello Barbie" smart toy, the Mattel doll equipped with software that records conversations around it and transmits them to a California company specializing in AI, with the aim of improving the relevance and quality of the interaction the toy has with its young owner. On the matter, see I.D. Manta, D.S. Olson, *Hello Barbie: First They Will Monitor You, Then They Will Discriminate Against You. Perfectly, 67 Alabama Law Review,* I (2015), p. 135 ff.

video games[8]. On the other hand, the minor represents at the same time the perfect victim of those same mechanisms.

The first concerns arise, of course, on the side of personal data, including that of minors, around which now well-known economic interests revolve, which have turned the data themselves almost into a necessary price to pay to access services, especially digital ones[9].

It is clear that the potential fragility of minors, especially when contextualized in a risky virtual world, demands special attention[10]. Indeed, the consequences of uncontrolled exposure to the risks of the digital ecosystem can become devastating with respect to subjects of whom, being physiologically in a psycho-physical condition of vulnerability, it is easy to influence their capacity for self-determination[11].

It is also true that the use of various digital tools by minors represents a form of manifestation of their personal and digital identity, integrating a decisive moment in the formation of their personality[12], in a context in which the physical world and the virtual world represent two articulations of the same space of relationship[13].

[8] On which see *infra*, para. 3.

[9] Without wishing here to retread the heated debate on the capitalization/contractualization of personal data, animated mainly by EU Directives 770 and 771 of 2019, later implemented in Italian legislation, we refer to extensive bibliography on the subject. See, *ex multis*, G. Alpa, *La "proprietà" dei dati personali»*, in N. Zorzi Galgano (ed.), *Persona e mercato dei dati. Riflessioni sul GDPR,* CEDAM, Padua, 2019, p. 18 ff., THOBANI S., Diritti della personalità e contratto. Dalle fattispecie più tradizionali al trattamento in massa dei dati personali, Turin, 2019, p. 158 ff., V. Ricciuto, *L'equivoco della privacy. Persona vs dato personale*, E.S.I., Naples, 2022, p. 105 ff., Id., *La patrimonializzazione dei dati personali. Contratto e mercato nella ricostruzione del fenomeno*, in V. Cuffaro,R. D'Orazio, V. Ricciuto, (eds.), *I dati personali nel diritto euopeo,* cit. 23 p. ff.,A. De Franceschi, *Il "pagamento" mediante dati personali, ivi,* p. 1381, Id., *La circolazione dei dati personali tra privacy e contratto*, E.S.I., Naples, 2017, p. 16, where it is highlighted that it is imperative to take into account the economic value of personal data and the correlative aspects of property law. In this regard, see also V. Janecek, *Ownership of personal data in the Internet of Things,* in «Computer Law & Security Review», V (2018), p. 1039 ff., F. Bravo, *Lo "scambio di dati personali" nei contratti di fornitura di servizi digitali e il consenso dell'interessato tra autorizzazione e contratto*, in «Contr. impr.», XXXV (2019), p. 44 ff.

[10] The delicate relationship between young users and the Internet is investigated, among others, by A. Mantelero, *Teens online and data protection in Europe*, in «Contr. impr. Europa», I (2014), p. 442 ff., Id., *Children online and the future EU data protection framework: empirical evidences and legal analysis*, in «Int. J. Technology Policy and Law», 2016, p. 169 ff., C. Perlingieri, L*a tutela dei minori di età nei social networks,* in «Rass. dir. civ», IV (2016), p. 1324 ff. With specific regard to the school environment see again A. Mantelero, *Adolescenti e privacy nella scuola ai tempi di YouTube*, in «Nuova giur. civ. comm.», II (2011), p. 139 ff.

[11] Talks about the «vulnerability» of the «electronic body» of «digital native minors», A. Astone, *I dati personali dei minori in rete. Dall'internet delle persone all'internet delle cose,* Milan, 2019, p. 5 ff. On this profile, see also I. Garaci, *Il "superiore interesse del minore" nel quadro di uno sviluppo sostenibile dell'ambiente digitale,* in «Nuova giur. civ. comm.», (2021), p. 801.

[12] See A. Thiene, *I diritti della personalità dei minori nello spazio virtuale,* in «Annali online did. e form.», vol.9, XIII (2017), p. 26 ff.

[13] See R. Caterina, *La libertà di comunicazione: il fenomeno dei social network*, in C. Perlingieri, L. Ruggeri (eds.), *Internet e Diritto civile*, E.S.I., Naples, 2015, p. 117 ff.

If the guarantee of constitutional principles, first and foremost Article 2 of the Constitution, implies that the minor can freely explicate and evolve his or her personality, it means that the minor can freely move in that direction for the realization of their identity interests[14].

2. Minors and Artificial Intelligence Act

It is known that the last crucial building block of the so-called European Data Strategy[15], is the Regulation of the European Parliament and of the Council laying down harmonized rules on Artificial Intelligence, the so-called Artificial Intelligence Act, AI Act, dated June 13th, 2024, aiming to improve the functioning of the internal market and promote the adoption of reliable and human-centered Artificial Intelligence, while ensuring a high level of protection of fundamental rights enshrined in the EU Charter[16]. The Regulation, effective August 1st, 2024, applies to public and private entities, inside and outside the EU, provided that the AI system is placed on the EU market or its use affects persons located in the EU[17].

It has already been mentioned about the impacts that AI systems can also and especially have on minor users. However, admittedly, in the AI Act, there are very few precipitous references to minors. Apart from general references to the protection of the fundamental rights of minors even in the digital environment, and the need not to harm them (see recitals 9, 28 and 48), there are no provisions specifically addressed to the minor.

Yet, certain aspects of the AI Act also appear to be quite interesting with respect to the minor user of AI systems.

[14] R. Senigaglia, *Minore età e contratto. Contributo alla teoria della capacità, Turin, 2020, p. 75 ff., ID, L'identità personale del minore di età nel cyberspazio tra autodeterminazione e parental control system*, in «Nuove leggi civ. comm.», no. 6, p. 1568 ff. And formerly see F.D. Busnelli, *Immagini vecchie e nuove della tutela della salute del minore,* in A. Bucelli (ed.), *Identità e salute del minore,* Pisa, 2021, p. 3. More recently see also the interesting reflections of D. Marcello, *Circolazione dei dati del minore tra autonomia e controllo. Norme e prassi nel mercato digitale europeo*, Naples, 2023, p. 51 ff.

[15] The European Data Strategy, which includes the "GDPR," EU reg. 2016/679, EU reg. 2018/1725 on non-personal data, the so-called Open Data Directive, to the Digital Service Act to the Digital Market Act, the Data Act, the Digital Governance Act, and the AI Act, predicts that the new rules will create "€270 billion of additional GDP for EU Member States by 2028 by addressing the legal, economic and technical issues that lead to data being underused." For an analysis of the strategy, see D. Poletti, *Il quadro normativo del Data Governance Act: l'esercizio dei diritti dell'interessato nell'attività di intermediazione dei dati*, in «Nuove leggi civ. comm.», III (2024), p. 533 ff.

[16] In this regard, see G. Finocchiaro, *Intelligenza artificiale. Quali regole?*, Bologna, 2024, p. 44. ff.

[17] It should be noted, however, that the application of the entire Regulation for all risk categories will be achieved only 36 months after the Regulation itself comes into force, which provides a kind of differentiated "roadmap."

First comes the consideration of Art. Art. 4, significantly titled "AI literacy," according to which AI system providers are not only required to take "measures to ensure, to their best extent, a sufficient level of AI literacy of their staff and other persons dealing with the operation and use of AI systems on their behalf," but they must do so while also taking into account "the persons or groups of persons on whom the AI systems are to be used."

Therefore, there will have to be, by the providers of AI systems, specific regard and even training parameterized also to the audience on which their system is intended to produce effects: one can therefore immediately assume the need to pay special attention to systems that are or may be intended for use by minors. In such a case, the required "literacy" should experience a decisive rise in level, consequent to the increased risks associated with the minor's age, and thus the (at least) potential vulnerability of the individuals in question.

First, it is noted how significantly a system based on risk management and containment comes back into play, in line with previous acts of the European Union, first among them the well-known GDPR[18], and then, with specific regard to minors, Article 9 is also taken into consideration, which states in para. 9, that when implementing the risk management system relating to high-risk AI systems, providers shall give consideration to the possibility that the high-risk AI system may have "an adverse impact on persons under the age of 18 and, as appropriate, other vulnerable groups".

Consistent with what is already specified in Article 4, special consideration must be given to the circumstance where AI systems are used by vulnerable individuals, including minors.

On the other hand, simplifying, the Regulation is strongly inspired by the concept of "risk" that now dominates the regulation of new technologies, precisely on the allocation of AI systems according to the risk they generate[19]. These range from unacceptable risk, which makes the use of the relevant systems prohibited, to high risk, such that the relevant AI systems are required to meet stringent requirements under the Regulations, including measures to mitigate that risk, to risk identified as minimal, with the result that the relevant AI systems are not subject to any obligation under the AI Act, but providers are encouraged to voluntarily adopt

[18] On which see, exemplarily, A. Mantelero, *Responsabilità e rischio nel Reg. Ue 2016/679,* in «Nuove leggi civ. comm.», I (2017), p. 144 ff.

[19] See G. Finocchiaro, *La regolazione dell'intelligenza artificiale*, in «Riv. trim dir. pubbl.»., IV (2022), p. 1085 ff., especially p. 1093 ff.

additional codes of conduct. There is also a specific "transparency risk": as clarified by the European Commission itself in a statement, systems that fall under it, such as chatbots, must clearly inform users that they are interacting with a machine, while some AI-generated content must be explicitly labeled as such[20].

3. An underestimated risk: the use of video games by minors

The category of the minimal risk of Artificial Intelligence systems, when compared to their use by minors, raises immediate perplexity, especially when the Commission, in its aforementioned communiqué, attributes to AI systems characterized by minimal risk, by way of example, "AI-enabled video games."

In this regard, it is even superfluous to point out how minors of age, are the users par excellence of video games, despite the fact that upstream there are actual doubts in this case about the validity of the relevant purchase/supply contract, depending on whether it is a physical purchase or whether it is performed through an online download.

Sometimes in fact users are anything but great minors capable of self-determination and of making autonomous and prudent choices in function of the development of their personality, assuming that such a function can be configured in relation to the purchase of such a product/service.

And even if it were a matter of giving consent to the processing of the user's personal data in the context of an information society service, often the minor-data subject is well below the age of fourteen. Moreover, there is no control so that the purchase is limited to video games indicated as suitable for the age group in which the minors themselves falls.

Again, it is rare that video games do not make use of solutions declined in terms of chatbots: the configuration of a customer service answering FAQs is sufficient for this purpose; and furthermore, in narrative video games, it is precisely a chatbot that appears by default, perhaps with human features, to answer the player's questions.

When video games integrate AI-powered conversational agents or other interactive AI functionalities, transparency obligations under the AI Act may become relevant. The legal assessment should therefore focus on the specific AI component and its intended use, rather than on the video game as a whole.

[20] See the official communiqué «The AI Regulation enters into force» published on the European Commission website, https://commission.europa.eu/ on August 1st, 2024.

But the dangers potentially posed to minors go far beyond interaction with a chatbot.

On the side of processing the personal data of young players, it is not always clear what data the video game collects and processes, and with the application of what security measures. If they were users located in the European Union, it is known that by virtue of the extraterritorial scope of the GDPR, all the relevant requirements would moreover have to apply[21].

A recent U.S. case that uncovered a real "Pandora's box" in the world of online gaming is exemplary in this sense: in 2022 Epic Games, the company that owns the famous game Fortnite, was the recipient of a very substantial fine following a settlement with the Federal Trade Commission[22].

The challenged conduct pertained precisely to the collection of personal data of users under the age of 13[23] — such as their names, email addresses, identifiers used to track players' progress, purchases made, game settings, and friends lists — data that the company that owns the processing knew it was collecting, the investigation found.

Another allegation that in turn led to the imposition of the sanction concerned the manipulation of minors put in place by Epic Games, which, in order to push minor users to purchase virtual goods used dark patterns, essentially carrying out real unfair business practices, inducing minor players to make purchases that could take place without parental consent[24].

Yet there is even more. In fact, the more alarming aspect arising from the use of video games by minors coincides with the possibility, offered by many video games, of interacting with other players, through messaging systems, or directly with conversations that take place through headsets equipped with a microphone.

[21] See A. Spangaro, *L'ambito di riferimento materiale del nuovo Regolamento,* in G. Finocchiaro (ed.), *Il nuovo Regolamento europeo sulla privacy e sulla protezione dei dati personali*, Bologna, 2017, p. 23 ff.

[22] On the case see S. Sarzana Di, F. Ippolito, *Fortnite viola la privacy di minori e li inganna: così la super sanzione da 520 milioni di dollari*, in «cybersecurity360.it», December 20th, 2022.

[23] The 13-year age limit derives from the Children's Online Privacy Protection Act (COPPA), an act passed by the U.S. Congress in 1998 and in effect since 2000, which dictates rules intended for website operators and online service providers aimed at protecting children's personal data and safety. What's more, the text of that law was updated in 2013 to reflect changes in technology, with the penalties for violating its requirements also tightened. In particular, when processing personal data of children under the age of 13, COPPA stipulates that it is the parents who must provide the relevant consent. On COPPA see S. Grandison, *The Child Online Privacy Protection Act: The Relationship Between Constitutional Rights and The Protection of Children*, in *University of the District of Columbia Law Review,* I (2011), p. 3-31-201.

[24] See again S. Sarzana Di, F. Ippolito, *Fortnite viola la privacy di minori e li inganna: così la super sanzione da 520 milioni di dollari*, cit.

[25] Article 28 DSA is devoted primarily to «Online protection of minors». It stipulates that provi-

Right from the start, it could be argued that offering the possibility for gamers to communicate with each other changes the nature of the video game by making it (also) a digital platform, consequently also subject to the obligations of the well-known Digital Services Act, EU Reg. 2022/2065, with its intertwining of applicable regulations. From the application of the DSA would derive, among other things, the prohibition on the profiling of minors[25], which is difficult to comply with where the video game records the conversations of minors and consequently proposes advertising targeted at them, intercepting their consumption needs, as occurred precisely in the Forntite case.

Moreover, advertising aimed at minors could further integrate a deceptive commercial practice, for that reason alone prohibited and sanctioned, as, again, established in the case involving the famous video game[26].

Attention thus returns to personal data collected as part of conversations, which could be recorded: it is no coincidence that online gaming sites are reporting the news, dated 2020, that Sony had announced that the soon-to-be-released Playstation 5 would record conversations between players for the purpose of moderating them within game groups[27].

Conversations, written or audio, even of underage players, could in fact be recorded or saved by many video games, without them being

ders of online platforms accessible to minors are required to «put in place appropriate and proportionate measures to ensure a high level of privacy, safety, and security of minors, on their service» (para. 1) they are prohibited from profiling minors of age (para. 2), albeit with not exactly crystal-clear wording. However, precisely in relation to the category of minors, targeted advertising, based on profiling, appears to be prohibited tout court: therefore, the prohibition of profiling in favor of minors seems to have finally been enshrined, which is not present, however, in Reg. 2016/679, even though the prohibition of processing minors' data for commercial purposes had, to be true, already been established by the EU Directive 2018/1808 on Audiovisual Media Services. The Digital Services Act also strengthens the protection of minors by limiting the use of profiling-based advertising practices directed at them, thereby reducing exposure to manipulative commercial strategies in online environments. See B. Agostinelli, *Minori in rete: l'illusione del consenso e l'equivoco della "privacy",* in «Familia», III (2023), p. 415. On this point see also G. Scorza, *Digital services act. Le luci e le poche ma gravi ombre delle nuove regole Ue*, in «agendadigitale.it», April 28th, 2022.

[26] The reference, as is well known, is to the Unfair Commercial Practices Directive 2005/29/EC, which is the general EU legislation governing unfair commercial practices that occur before, during and after transactions between professionals and consumers. In Italy, it was transposed by Legislative Decree no. 146/2007, which interpolated the Consumer Code with Articles 18 to 27. On the topic, see, *ex multis*, G. De Cristofaro, «La direttiva 2005/29/CE, in Id (ed), Pratiche commerciali scorretta e codice del consumo», Giappichelli, Turin, 2008, p. 1 ff.

[27] See the news that came out on October 17th, 2020 at https://gaming.hwupgrade.it/, and the official statement, dated October 16th, 2020, from Catherine Jensen, president of the Global Consumer Experience division https://blog.playstation.com/2020/10/16/details-on-new-voice-chat-functionality-coming-to-ps5/, who does not deny the activation of the feature in question and, indeed, even points out the technical impossibility of disabling it, while generically declaring it to be in compliance with privacy regulations.

provided with full disclosure, clarifying, for example, how they would be recorded and the retention times, as well as specific, compliant with the more stringent requirements already set forth in the GDPR.

The additional passage in the outlined scenario is, of course, the interconnection of the video game with headphones to the other smart devices in the child's home, primarily home speakers or smart assistants, as part of a home automation environment pervaded by the operation of the Internet of Things[28].

Moreover, not least, the serious dangers of grooming that arise for the minor, especially the infant, in interfacing with other users, even adults, who may come into possession of important information about the minor himself or herself, as well as the dangers of becoming a victim of conduct qualifying in terms of cyberbullying, seem obvious.

In this regard, therefore, Legislative Decree No. 39/2014, implementing EU Directive 2011/93 on combating the sexual abuse and sexual exploitation of minors and child pornography, recalled by the DSA itself[29], should come into play, also in conjunction, on the one hand, and Law No. 71 of 2017, aimed at preventing and combating the phenomenon specifically of cyberbullying, as well as the very recent Law No. 70/2024, also bearing provisions aimed at preventing and opposing the phenomena of both bullying and cyberbullying[30].

Without wishing to delve into the albeit relevant topic of cyberbullying here, and in an attempt to draw some initial summary conclusions, it is noted how the AI Act decries in several places the limitation to the operation of regulated technologies represented by respect for the fundamental rights of the child. The AI Act even cites Article 24 of the Charter of fundamental rights of the European Union, which is best known for enshrining the right of the child to express his or her opinion and to be heard; in fact, this provision is far more

[28] See the Italian Data Protection Authority information sheet released in March 2020, on Digital assistants (smart assistants): the Authority's advice for privacy-proof use, which highlights how smart assistants can be easily interconnected with other objects and services in home spaces, such as home appliances, smart TVs, installations, and even video games. In this way, the presence of a single smart product facilitates the performance of home automation functions by other devices as well.

[29] Regulation 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market for Digital Services and amending Directive 2000/31/EC (Digital Services Act), entered into force on February 17th, 2024.

[30] On combating the phenomenon of cyberbullying, see C. Perlingieri, «Profili civilistici dei social network», Naples, 2014, p. 33 ff., and more recently, E. Battelli, *Minori e social network: cyberbullismo e limiti della parental responsibility*, in «Corr. giur.», X (2021), p. 1269 ff., E. Andreola, *Minori e incapaci in Internet,* cit., p. 243 ff., R. Senigaglia, *Il dovere di educare i figli nell'era digitale*, in «Pers. merc.», III (2021), cit., p. 521 ff., F. Zanovello, *Prevenzione e contrasto del bullismo e del cyberbullismo. Tra novità e criticità della l. n. 70/24*, in «Nuove leggi civ. comm.», IV (2024), p. 826 ff.

far-reaching in its view of the overall protection of the child person, where it opens by stating that "Children shall have the right to such protection and care as is necessary for their well-being."

The well-being, precisely, of minors, the now famous "best interest of the child"[31] should, if not inspire the latest European regulations on personal data and Artificial Intelligence, at least put an effective brake on the expansionist ambitions of the Big Players who dominate the relevant markets, and as noted, although in a network marked by the presence of a plurality of players, it is precisely these latter that dictate the rules; it is no coincidence that the very same States of law[32] dialogue with them, or rather "bargain" with them.

However, the limitation of child well-being seems to remain confined to a declaration of intent, where reality shows that the pitfalls behind children's use of new technologies often escape the meshes of a regulation that is also overabundant, detailed, but evidently fails to grasp the complexities of the phenomenon, especially from the point of view of the most vulnerable.

4. A first conclusion: the role of certifications and Cybersecurity

In this context, Cybersecurity certifications may represent a practical bridge between regulatory requirements and technological implementation. They can provide verifiable standards of protection for products and services frequently used by minors, promoting trust, accountability and greater consistency across the digital market.

[31] The literature on the subject is vast. *Ex multis*, and also with different perspectives, see E. Quadri, *L'interesse del minore nel sistema della legge civile*, in «Fam. dir.», (1999), p. 80 ff., L. Lenti, "*BEST interests of the child" o "best interests of children"?*, in «Nuova giur. civ. comm.», (2010), p. 157 ff., Id., *Note critiche in tema di interesse del minore*, in «Riv. dir. civ.», (2016), p. 86 ff, V. Scalisi, *Il superiore interesse del minore, ovvero il fatto come diritto,* in «Riv. dir. civ.», II (2018), p. 405 ff., M. Sesta, *La prospettiva paidocentrica quale fil rouge dell'attuale disciplina giuridica della famiglia*, in «Fam. dir.», VII (2021), p. 763 ff., E. Bivona, *Diritto vigente e diritto vivente nell'affidamento dei figli,* in «Familia», VI (2022), p. 803 ff., E. Lamarque, *Pesare le parole. Il principio dei best interests of the child come principio del miglior interesse del minore,* in Fam. dir., VI (2023), p. 365 ff., Id., *Prima di bambini. Il principio dei best interests of the child nella prospettiva costituzionale,* Milano, 2016, passim, especially p. 64 ff., G. Recinto, *Il superiore interesse del minore tra prospettive interne "adultocentriche" e scelte apparentemente "minoricentriche" della Corte europea dei diritti dell'uomo*, in «Foro it,», I (2017), 3669 ff., M. Dogliotti, *Protezione dei minori, diritto interno, giurisprudenza della Corte eur. dir. uomo,* in «Nuova giur civ. comm.», I (2023), p. 420 ff.

[32] See the analysis of D. Poletti, *A proposito di fonti nell'"ecosistema digitale",* in F. Ricci (ed), Principi, clausole generali, argomentazione e fonti del diritto, Milan, 2018, p. 348 ff. The related risks in terms of injury to personal dignity have been well highlighted by S. Rodotà, *Il mondo nella rete. Quali i diritti, quali i vincoli*, Roma-Bari, 2014, p. 27 ff.

Against this background, preventive mechanisms such as certification schemes contribute to the identification, mitigation and management of risks affecting vulnerable users, including minors.

The certification system is certainly not a novelty in the European panorama, starting with the consolidated certification apparatus already in place, relating to product safety.

In fact, through the affixing of the CE marking — whose rules are established by Article 30 of Regulation 765/2008 — the manufacturer certifies the conformity of its product with the essential safety requirements laid down in one or more Community Directives applicable to it[33].

Similarly, a European scheme for Cybersecurity certification of ICT products and digital services was prepared with the Cybersecurity Act of December 2018[34].

Moreover, in 2023, the European Commission has proposed an amendment to the Cybersecurity Act aimed precisely at the future adoption of European certification schemes albeit limited to particularly sensitive areas[35].

Precisely with regard to the field of Internet of Things products, such as video games, in particular, the issue of interoperability between different existing certification mechanisms will therefore arise. It will therefore be a matter of understanding how the different disciplines inherent in certifications involving the same product — in data protection on the one hand, Cybersecurity on the other, but also product security more generally — will integrate.

The questions that arise are many and varied. To what extent can different certifications insist on the same product? What value does a certification inherent in a certain sector have with regard to demonstrating compliance with a different discipline? Can the presence of a certain certification exempt one from obtaining another?

[33] In the General Product Safety Directive 2001/95/EC, a product is considered safe if, under normal or reasonably foreseeable conditions of use, including durability and, where applicable, putting into service, installation and maintenance, it presents no risk or only minimal risks that are compatible with the product's use and are considered acceptable in observance of a high level of protection for the health and safety of persons. Article 4 of the Machinery Directive 2006/42/EC requires member states to ensure that the provisions relating to machinery and partly completed machinery are properly implemented and that machinery placed on the market and put into service is safe.

[34] See D. Poletti, M.C. Causarano, *Autoregolamentazione privata e tutela dei dati personali: tra codici di condotta e meccanismi di certificazione*, in E. Tosi (ed), *Privacy digitale, persona e mercato: tutela della riservatezza e protezione dei dati personali alla luce del GDPR e del nuovo Codice Privacy,* Milan, 2019, p. 369 ff.

[35] See the EC's own statement at https://digitalstrategy.ec.europa.eu/en/policies/cybersecurity-act

Still, a further development-which, however, at present represents a mere supposition-could follow the direction of Quality Management Systems (UNI EN ISO 9001:2015).

Indeed, corporate structures already equipped with such a system could also integrate internally personal data processing procedures and/or Cybersecurity techniques[36] even in parallel with the attainment of possible certifications.

The outcome could be an improvement in the effectiveness and efficiency of product realization or service delivery, including from the point of view of compliance with the relevant applicable regulations.

In any case, the anticipation of the precautions to be taken by the manufacturer already at an entirely preliminary stage, of design, of product design, is, moreover, entirely in keeping with precisely that principle of privacy by design that already permeates the GDPR regulations.

In this way, it should be achieved to put on the market a product that is safe both "physically" (e.g., built with materials that are not harmful to the user's health) and from the point of view of the processing of personal data that it carries out, which will have to comply with the principles of data protection, and, more generally, in its operating processes. It being understood that the need to anticipate damage, to prevent it with the techniques of security by design, does not obviate the need to properly identify liability profiles, when damage occurs anyway.

The ultimate goal should be to achieve an integrated system of overall security of the product or service, with a view to the achievement of an equally integrated protection of the individual, even the minor, against a plurality of related risks.

A preventive approach based on digital literacy, Cybersecurity by design and independent certification mechanisms may contribute to creating a safer digital ecosystem for younger generations, while preserving innovation and fundamental rights.

[36] In particular, incident response approaches. Useful may be the ISO 31000:2018 standard on risk management, which aims to put every organization in a position to identify, prevent and manage all impending risks within its business through a structured approach.

BIBLIOGRAPHY

Agostinelli B, *Minori in rete: l'illusione del consenso e l'equivoco della "privacy"*, in «Familia», III (2023), p. 397 ff.

Alpa G., *La "proprietà" dei dati personali*, in N. Zorzi Galgano (ed.), *Persona e mercato dei dati. Riflessioni sul GDPR*, Padua, 2019, p. 18 ff.

Alpa G., *Tecnologie e diritto privato*, in «Riv. it. sc. giur.» (2017), p. 205 ff.

Andreola E., *Minori e incapaci in Internet*, E.S.I., Naples, 2019.

Astone A., *I dati personali dei minori in rete. Dall'internet delle persone all'internet delle cose*, Giuffrè, Milan, 2019.

Battelli E., *Minori e social network: cyberbullismo e limiti della parental responsibility*, in «Corr. giur.», X (2021) p. 1269 ff.

Bivona E., *Diritto vigente e diritto vivente nell'affidamento dei figli*, in «Familia», VI (2022), p. 803 ff.

Bravo F., *Lo "scambio di dati personali" nei contratti di fornitura di servizi digitali e il consenso dell'interessato tra autorizzazione e contratto*, in «Contr. impr.», XXXV (2019), p. 44 ff.

Busnelli F.D., *Immagini vecchie e nuove della tutela della salute del minore*, in Bucelli A. (ed.), *Identità e salute del minore*, Pacini, Pisa, 2021, p. 3 ff.

Caterina R., *La libertà di comunicazione: il fenomeno dei social network*, in Perlingieri

C., Ruggeri L. (eds.), *Internet e Diritto civile*, E.S.I., Naples, 2015, p. 117 ff.

Di Ciommo F., *Evoluzione tecnologica e regole di responsabilità civile*, Naples, 2003.

De Cristofaro G., *La direttiva 2005/29/CE*, in Id (ed), *Pratiche commerciali scorretta e codice del consumo*, Turin, 2008, p. 1 ff.

De Franceschi A., *Il "pagamento" mediante dati personali*, in Cuffaro V., D'Orazio R.,

Ricciuto V., (eds.), *I dati personali nel diritto europeo*, Milan, 2019, p. 1381 ff.

Di Landro A.C., *Best interest of the child e tutela dei minori nel dialogo tra legislazione e giurisprudenza*, in «Nuove leggi civ. comm.», II (2020)

Dogliotti M., *Protezione dei minori, diritto interno, giurisprudenza della Corte eur. dir. uomo*, in «Nuova giur civ. comm.», I (2023), p. 420 ff

Finocchiaro G., *Intelligenza artificiale. Quali regole?*, Bologna, 2024.

Finocchiaro G., *La regolazione dell'intelligenza artificiale*, in «Riv. trim dir. pubbl.»., IV (2022), p. 1085 ff

Garaci I., *Il "superiore interesse del minore" nel quadro di uno sviluppo sostenibile dell'ambiente digitale*, in «Nuova giur. civ. comm.», (2021), p. 801.

Giovanella F., *Le persone e le cose: la tutela dei dati personali nell'ambito dell'Internet of Things*, in in Cuffaro V., D'Orazio R., Ricciuto V., (eds.), *I dati personali nel diritto europeo*, Milan, 2019, p. 1213 ff.

Janeček V., *Ownership of personal data in the Internet of Things*, in «Computer Law & Security Review», V (2018), p. 1039 ff.

Lamarque E., *Pesare le parole. Il principio dei best interests of the child come principio del*

miglior interesse del minore, in «Fam. dir.», IV (2023), p. 365 ff.

Lamarque E., *Prima di bambini. Il principio dei best interests of the child nella prospettiva costituzionale*, Milano, 2016.

Lenti L., *"Best interests of the child" o "best interests of children"?,* in «Nuova giur. civ. comm.», (2010), p. 157 ff.

Lenti L., *Note critiche in tema di interesse del minore*, in «Riv. dir. civ», 2016, p. 86 ff.

Manta I.D, Olson D.S., *Hello Barbie: First They Will Monitor You, Then They Will Discriminate Against You. Perfectly*, 67 *Alabama Law Review*, I (2015), p. 135 ff.

Mantelero A., *Adolescenti e privacy nella scuola ai tempi di YouTube*, in «Nuova giur. civ. comm.», II (2011), p. 139 ff.

Mantelero A., *Children online and the future EU data protection framework: empirical evidences and legal analysis*, in «Int. J. Technology Policy and Law», (2016), p. 169 ff.

Mantelero A., *Responsabilità e rischio nel Reg. Ue 2016/679*, in «Nuove leggi civ. comm.», I (2017), p. 144 ff.

Mantelero A., *Teens online and data protection in Europe*, in «*Contr. impr. Europa*», I (2014), p. 442 ff.

Marcello D., *Circolazione dei dati del minore tra autonomia e controllo. Norme e prassi nel mercato digitale europeo*, E.S.I., Naples, 2023.

Nardi S., *La famiglia e gli affetti nell'era digitale*, Naples, 2020.

Oppo G., *Scienza, diritto, vita umana*, in «Riv. dir. civ.»,I (2002)

Perlingieri C., *La tutela dei minori di età nei social networks*, in «Rass. dir. civ.», IV (2016), p. 1324 ff.

Perlingieri C., *Profili civilistici dei social networks*, E.S.I., Naples, 2014.

Poletti D., *A proposito di fonti nell'"ecosistema digitale"*, in Ricci F. (ed), *Principi, clausole generali, argomentazione e fonti del diritto*, Giuffrè, Milan, 2018, p. 348 ff.

Poletti D., Causarano M.C., *Autoregolamentazione privata e tutela dei dati personali: tra codici di condotta e meccanismi di certificazione*, in Tosi E. (ed), *Privacy digitale, persona e mercato: tutela della riservatezza e protezione dei dati personali alla luce del GDPR e del nuovo Codice Privacy*, Giuffrè, Milan, 2019, p. 369 ff.

Poletti D., *Il quadro normativo del Data Governance Act: l'esercizio dei diritti dell'interessato nell'attività di intermediazione dei dati*, in «Nuove leggi civ. comm.», III (2024), p. 533 ff.

Quadri E., *L'interesse del minore nel sistema della legge civile*, in «Fam. dir.», 1999, p. 80 ff.

Ramaccioni G., *La protezione dei dati personali e il danno non patrimoniale. Studio sulla tutela della persona nella prospettiva risarcitoria*, Naples, 2017.

Recinto G., *Il superiore interesse del minore tra prospettive interne "adultocentriche" e scelte apparentemente "minoricentriche" della Corte europea dei diritti dell'uomo*, in «*Foro it.*», 2017, I, c. 3669 ff.

Ricciuto V., *L'equivoco della privacy. Persona vs dato personale*, E.S.I., Naples, 2022.

Ricciuto V., *La circolazione dei dati personali tra privacy e contratto*, E.S.I., Naples, 2017

Ricciuto V., *La patrimonializzazione dei dati personali. Contratto e mercato nella ricostruzione del fenomeno*, in Cuffaro V., D'Orazio R., Ricciuto V., (eds.), *I dati personali nel diritto euopeo*, Giappichelli, Torino, p. ff.

Rodotà S., *Il mondo nella rete. Quali i diritti, quali i vincoli*, Roma-Bari, 2014, p. 27 ff.

Santosuosso A., *Intelligenza artificiale e diritto*, Milan, 2020.

Sarzana di S. Ippolito F., *Fortnite viola la privacy di minori e li inganna: così la super sanzione da 520 milioni di dollari*, in «cybersecurity360.it», December 20th, 2022.

Scalisi V., *Il superiore interesse del minore, ovvero il fatto come diritto*, in «Riv. dir. civ.», 2018, no. 2, p. 405 ff.

Scorza G., *Digital services act. Le luci e le poche ma gravi ombre delle nuove regole Ue*, in «agendadigitale.it», April 28th, 2022.

Senigaglia R., *Il dovere di educare i figli nell'era digitale*, in «Pers. merc.», 2021, no. 3 cit., p. 521 ff.

Senigaglia R., *L'identità personale del minore di età nel cyberspazio tra autodeterminazione e parental control system*, in «Nuove leggi civ. comm.», no. 6, p. 1568 ff.

Senigaglia R., *Minore età e contratto. Contributo alla teoria della capacità*, Turin, 2020.

Sesta M., *La prospettiva paidocentrica quale fil rouge dell'attuale disciplina giuridica della famiglia*, in «Fam. dir.», 2021, no 7, p. 763 ff.

Spangaro A., *L'ambito di riferimento materiale del nuovo Regolamento*, in Finocchiaro G. (ed.), *Il nuovo Regolamento europeo sulla privacy e sulla protezione dei dati personali*, Bologna, 2017, p. 23 ff.

Thobani S., *Diritti della personalità e contratto. Dalle fattispecie più tradizionali al trattamento in massa dei dati personali*, Turin, 2019.

Thiene A., *I diritti della personalità dei minori nello spazio virtuale*, in «Annali online did. e form.» doc., Vol. 9, No. 13/2017, p. 26 ff.

Vizzoni L., *Domotica e diritto*. La *Smart Home tra regole e responsabilità*, Milan, 2021, p. 72 ff.

Weber R. H., *Internet of Things-Need for a New Legal Environment?*, in «Computer Law & Security Review», 2009, p. 521 ff.

Zanovello F., *Prevenzione e contrasto del bullismo e del cyberbullismo. Tra novità e criticità della l. n. 70/24*, in «Nuove leggi civ. comm.», 2024, no 4, p. 826 ff.

Finito di stampare a Napoli nel 2026
presso l'Area University Press
dell'Università degli Studi Suor Orsola Benincasa