Pioneering environmental protection through data intelligence

SIEMENS

Executive summary

Tackling the global triple crisis of climate change, biodiversity loss, and environmental pollution demands a technologically advanced response. Regulatory mandates such as the Corporate Sustainability Reporting Directive (CSRD) require corporations to deliver unprecedented transparency and accountability with speed and accuracy that manual processes increasingly struggle to support. Ensuring environmental regulatory compliance in over 200 countries is both a critical challenge and a strategic opportunity for Siemens.

This paper demonstrates how we leverage Artificial Intelligence (AI) and advanced data analytics as a critical enabler for environmental sustainability. By establishing a specialized data analytics team focused on environmental protection that fueses deep environmental expertise with cutting-edge data infrastructure and AI capabilities, we are translating environmental protection into robust, data-driven solutions.

The impact is evident in our own operations. For example, the Siemens Environmental Reporting Assistant (SERA) automates water risk assessments, significantly reducing evaluation time while maintaining human oversight. Also, the Siemens Environmental Knowledge Graph (SiEKG) integrates fragmented hazardous substance data through intelligent and reliable estimations, enabling comprehensive reporting despite fragmented and heterogeneous source data. These AI-enabled solutions deliver reliable precision in environmental accounting, efficiency gains across processes, and accessible expertise across the organization.

By integrating environmental expertise with data intelligence, we can effectively address regulatory demands while setting new benchmarks in environmental management and sustainability reporting.

Contents

1 Why AI matters for environmental protection

Tackling the global triple crisis of climate change, biodiversity loss, and environmental pollution demands a concerted and technologically advanced response. We are at a rare juncture where technological capability, economic incentives, and environmental necessity converge, offering an unprecedented opportunity for a digital and sustainability transformation. Artificial Intelligence (AI) is a pivotal enabler of environmental sustainability in this era of fundamental change.

A turning point for environmental sustainability

The imperative for Environmental, Social, and Governance (ESG) transparency driven by regulatory and stakeholder pressure heightens the need for robust data and advanced technology to make sense of it. Where past approaches relied on spreadsheets and manual assessments, AI-enabled analytics transforms how data exploration and analysis are operationalized at scale. Data analytics uses advanced methods to collect, analyze, and visualize data to turn complex information into actionable insights – with AI-powered methods like machine learning, natural language processing, and generative AI further enhancing automation and predictive capabilities. These enhanced capabilities make it possible to process vast data volumes, identify complex patterns imperceptible to human analysis, improve data-driven decision-making, and fulfill regulatory requirements significantly faster and more accurately than manual processes.

Siemens' approach to data-driven environmental protection

Siemens continues to be at the forefront of this technological evolution. Our ONE Tech approach strategically positions us to harness the full potential of data and AI. We recognize that AI is a critical driver for scaling environmental protection, driving innovation, and empowering shared capabilities to deliver faster insights and greater operational efficiency across our operations.

„The complexity of juggling global interconnected system transitions in times of major disruptions can only be mastered with AI."
Dr. Eva Riesenhuber, Global Head of Sustainability at Siemens.

Siemens established a data analytics team for environmental protection to effectively advance environmental sustainability through data-driven solutions. This specialized team operates on the principle that effective environmental protection requires the synergistic fusion of deep environmental expertise and cutting-edge data infrastructure, data science, and AI capabilities. Uniquely positioned to bridge this critical gap, they translate complex environmental challenges into robust, data-driven solutions. This paper illustrates the strategic importance of fusing environmental expertise, advanced data, and AI capabilities. These pages present use cases that exemplify our commitment to advancing environmental sustainability by driving transparency, setting ambitious targets, and working towards achieving them.

2 Environmental data analytics and artificial intelligence: A synergistic approach

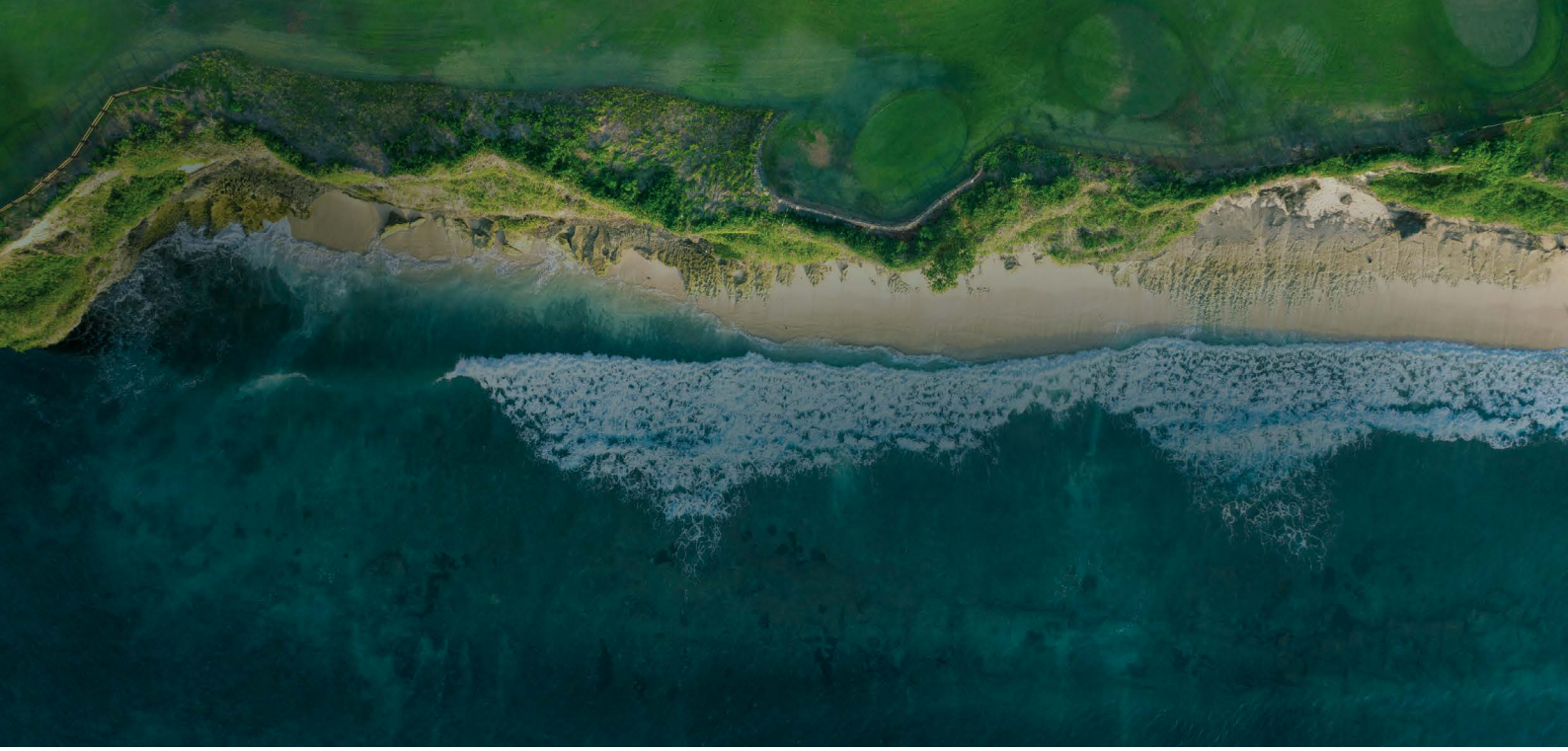
Regulatory momentum has increased scrutiny on environmental data, making its quality and defensibility decisive for compliance and credibility. Siemens strategically anchors data and AI within environmental protection, establishes traceable environmental data flows, and leverages data analytics and AI under rigorous human oversight. The result is a shift from manual, fragmented processes to a scalable, audit-ready practice with standardized assessments, faster KPI calculation, and earlier signals of operational and compliance risks.

Data-driven environmental stewardship

Environmental protection entered a new era in 2019/2020 as governments around the world unveiled ambitious "Green Deals". Globally, regulations increasingly demand greater transparency across critical environmental metrics to elevate environmental accountability – from pollution to nature impacts to circularity. The Corporate Sustainability Reporting Directive (CSRD) in the EU is one example of this move towards more rigorous sustainability reporting. This external momentum for environmental transparency only continues to accelerate.

"Legislators define what has to happen. But there is freedom on how to make it happen: This is where the innovation begins.

Klaus Luetzenkirchen,
Vice President Environmental Protection at Siemens, recognized the enormous potential of this momentum early on.



As a technology leader, Siemens is uniquely positioned to leverage AI to scale sustainability impact. We have set ambitious environmental targets to ensure meaningful progress, turning regulatory challenges into catalysts for AI-enabled environmental protection.

"We want to drive change. We want to accelerate the sustainability transformation – for our customers and for ourselves. And AI is a great means to do that."

Jochen Gross, Head of Sustainability Data Strategy and Architecture at Siemens.

These environmental targets are outlined in our CSRD-aligned Sustainability Statement and embedded within our DEGREE framework, which guides our sustainability priorities across our operations and portfolio. These targets include advancing the Siemens climate, zero pollution, biodiversity, and water strategies. We are committed to measurable progress on sustainability, using data analytics and AI-enabled environmental innovations to turn ambition into impact.

Integrating data analytics and environmental sustainability

Siemens established a strategic team to translate the synergistic commitment to environmental sustainability and data science into reality. This team is focused on environmental protection, reflecting the understanding that effective environmental responsibility relies on integrating environmental expertise with advanced data infrastructure, AI, and data science capabilities.

This setup effectively bridges the critical gap between expert environmental knowledge and cutting-edge technological approaches, such as AI and data science. It acknowledges that while IT departments excel in technical infrastructure, they often lack expertise in the ever-growing complexity of environmental protection requirements. Conversely, environmental experts may depend on data science expertise to fully harness AI's potential, identify relevant use cases, and establish the technical prerequisites for successful implementation.

This comprehensive integration helps data-driven environmental solutions be both technically sound and contextually relevant. In close collaboration, the team articulates real-world problems and implements and validates AI solutions. This approach ensures that developed tools are relevant, accurate, and seamlessly integrated into existing workflows.

By institutionalizing this expertise, we established the structural foundation needed to strategically leverage data science and AI for environmental protection.

Building the robust data foundation
Robust data is the superpower behind AI. High-quality data empowers AI to deliver accurate, reliable, and actionable insights – especially in industrial settings where decisions must reflect the real world.

"Building a strong data foundation was paramount for transforming environmental protection from a spreadsheet-based, largely manual operation into a cloud-native, AI-powered one. This required a new approach and architecture for environmental data management."

Kaiyuan Xin oversees data engineering, machine learning architecture, and cloud infrastructure at Siemens, with experience from environmental and supply chain analytics.

Historically, environmental data management operated in a landscape of decentralized spreadsheets, manual calculations, and email-based data exchanges. The lack of traceability, access controls, and scalability for advanced analytics created significant barriers, demanding a fundamental infrastructural overhaul for effective environmental data management.

The strategic response involved an end-to-end data cloud solution. Leveraging the corporate data infrastructure, environmental data is integrated and combined with financial and supply chain information. This approach ensures that environmental data, regardless of its origin, is centralized, accessible, and usable.

The challenge in unifying diverse data sources lies in the heterogeneity and variability of raw data. To address this, the team built data pipelines to cleanse, standardize, and transform raw data into unified figures, including semantic models. These are ready for use in AI models and effective processing with analytics tools.

The impact of this robust data foundation was visible immediately. Data processing, once requiring significant manual effort, including extensive method and formula checks by expert teams, is now automated and cloud-based. Streamlined data processing and automated KPI calculations have transformed company-level reporting, making it faster, more efficient, and more reliable.

The team operates within the broader corporate data infrastructure. This collaborative model ensures that the team leverages the enterprise data cloud, integrates seamlessly with existing identity and access management systems, and adheres to stringent security and governance standards. By designing solutions with open interfaces, the team actively promotes reusability across the organization. This integrated approach accelerates development cycles as the platform matures, helping AI-powered environmental solutions integrate seamlessly and effectively into native workflows and contributing to the organization's overall digital transformation.

Unlocking environmental AI

The initial focus on basic digitization – moving from spreadsheets to systematic data flows – paved the way for AI applications. The robust data foundation helps the team translate complex environmental challenges into innovative solutions using methods like machine learning and generative AI.

"Building effective AI solutions isn't just about algorithms. It's about rigorous human oversight and impeccable data governance."

Sven Kristen, Team Lead at Siemens, established the team and brings strategic vision and public sector experience in translating legal requirements into automated data processes.

Environmental data science experts must train, validate, and monitor models to ensure predictions align with regulatory context and operational reality. Moreover, poor data quality fundamentally limits the effectiveness of AI as systematic errors, biases, or gaps propagate through models. The team invests a great deal of effort in data validation, cleansing, model selection, and documentation to get reliable outputs, applying rigorous quality gates throughout the process.

Initial implementations demonstrated AI's immediate value. Namely, generative AI drafts standardize CSRD report sections, significantly reducing manual effort and ensuring consistency. Automated data merges consolidate information from multiple systems, eliminating manual work that once took days.

Every AI project is grounded in real-world environmental problems, developing solutions that directly address operational needs and contribute to strategic objectives.

"We didn't just want to build models. We wanted to build impact. Every project focuses relentlessly on genuine business problems, not technology showcases. The guiding question is always: Does this make someone's job easier, or decisions better?"

Sven Kristen

The solutions unburden Siemens business units, allowing a shift of focus from operational data processing to using analytical results effectively.

"We lverage machine learning to process messy, inconsistent data and find patterns that humans would miss. Generative AI allows us to synthesize information from dozens of sources into coherent, standardized assessments in minutes rather than hours. And, with advanced analytics, we can spot anomalies and emerging risks before they become problems. This is proactive, not reactive environmental management."

Sven Kristen

3 AI-enabled environmental data intelligence for reliable and scalable regulatory reporting

The strategic value of environmental data intelligence goes far beyond solving isolated operational challenges. Its impact lies in enabling scalable, reliable, and audit-ready environmental reporting across a rapidly evolving regulatory landscape. At Siemens, this means transforming complex environmental regulatory challenges into consistent, traceable, and disclosure-ready data flows.

This chapter highlights two capabilities that demonstrate this transformation in our own operations: the Siemens Environmental Reporting Assistant (SERA) for water risk assessments and the Siemens Environmental Knowledge Graph (SiEKG) for the estimation of hazardous substances. Both solutions address core reporting requirements – automating data synthesis to accelerate expert-driven assessments and closing disclosure-critical data gaps. Together, they show how AI-supported intelligence strengthens the full reporting chain: from structured data generation and quality assurance to audit-ready outputs.

Use case 1: Smart Environmental Reporting Assistant for water risk assessment

Integrating easily accessible AI-powered solutions into existing systems and tools is critical to effectively facilitate complex operations and enable 24/7 support for environmental protection assessment and consultation. A pioneering project in this area is the integration of an AI assistant into the Siemens Environmental Reporting System for water risk assessment, which supports our water strategy and tracks progress on defined targets for water resource preservation, reduction of town water withdrawal, and the mitigation of associated risks. Water risk assessments systematically evaluate the impact of operations on water resources. These assessments help identify vulnerabilities and guide mitigation strategies to ensure sustainable water management.

The need to overcome challenges in the existing water risk assessment process, particularly concerning scalability and consistency, directly drove the development of the Smart Environmental Reporting Assistant (SERA). Previously, environmental analysts could not request immediate support for these complex assessments, potentially leading to interruptions in the process and additional effort to seek guidance from corporate experts. While reliance on expert input helped ensure high quality, the demand for periodic reassessments across approximately 200 global facilities created temporary capacity constraints within the process. Standardization also proved difficult as different analysts often emphasized distinct information sources or interpreted findings divergently. This resulted in inconsistencies which had to be harmonized in time-consuming joint sessions with corporate experts. The reliance on individual expert knowledge created bottlenecks that traditional automation could not address. SERA was designed to tackle these limitations by leveraging AI to enhance process efficiency.

"The AI assistant was conceived not merely as a chatbot with simple Q&A functionality. Rather, we built it as a tool to provide guidance, helping navigate and summarize validated data, perform plausibility checks, and point potential discrepancies for expert to review."

Christian Gilabert Alarcón bridges environmental science with data analytics at Siemens and contributes to AI assistants, serving as a translator between environmental experts and technical implementations.

The approach centers on leveraging the strengths of Large Language Models (LLMs): synthesizing information from diverse sources while maintaining human expertise for interpretation. From a user's perspective, SERA operates seamlessly. Based on site-specific information, it derives relevant response considerations and scoring logic to support expert assessments.

A sophisticated architecture orchestrates multiple data components behind this streamlined interface. The system automatically retrieves data from multiple authoritative sources: water stress indicators from the World Resources Institute, and historic and projected climate data for the region.

It processes this information through a structured pipeline and carefully engineered prompts that instruct the language model to analyze from an environmental risk perspective, following the same approach that experienced analysts use.

Crucially, SERA embodies the principle of augmented intelligence. On demand, it provides experts with relevant information, contextual explanations, and response considerations based on facility-specific inputs.
SERA synthesizes available data and highlights potential discrepancies or areas that may require closer attention. While the AI handles time-consuming research and initial data synthesis upon request, experts remain responsible for validation and final approval, ensuring reliable risk assessments.

From a development perspective, the technical implementation required months of focused work. This involved building the underlying systems and interface, refining prompt engineering through systematic testing with real facility data, integrating multiple validation layers, and creating feedback loops with domain experts. Through this rigorous cycle, the team gained profound insights into the types of questions and data complexities AI could reliably handle, and where it required more explicit guidance or constraints. The result: an AI assistant that empowers environmental experts to identify data-informed mitigation measures.

The global roll-out of SERA yielded highly valuable benefits. The tool offers significant educational value, accelerating the learning curve for junior experts with its guided processes. It also helped senior experts reallocate substantial time previously allocated to routine assessments, allowing them to focus on more complex cases and strategic oversight. Building on this success, the team is extending this validated methodology to other critical risk assessment domains, including biodiversity impacts and regulatory compliance monitoring.

Use cwase 2: Siemens Environmental Knowledge Graph for estimation of hazardous substances

Managing hazardous substances is a critical and complex responsibility for industrial operations, directly impacting environmental compliance, supply chain integrity, and product safety. With thousands of substances subject to intricate and evolving regulatory classifications, these regulations enable companies like Siemens to transparently report on their use while ensuring environmental compliance.

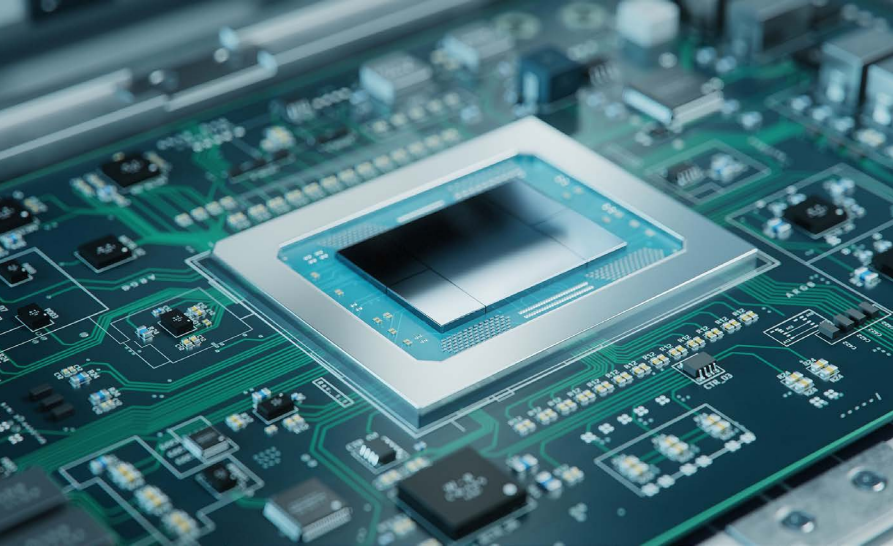
"SiEKG represents a fundamental shift from reactive, manual data handling to proactive, ontology-powered data representation that seamlessly integrates fragmented sources. It is specifically designed to enable complex environmental assessments and KPI calculations in real-world data environments."

Hansi Senaratne, responsible for setting up SiEKG at Siemens, combines computer science expertise with academic research in AI applications to solve complex environmental data challenges.

The core challenge in managing substance data at scale lies in the heterogeneous nature of this data distributed across numerous disparate sources. Critical information originates from a multitude of internal and external systems – product lifecycle management, procurement platforms, safety data sheets, and regulatory databases – each using different formats, naming conventions, and data structures. On top of this, CSRD requirements introduce the tracking of substance weights – information not previously required by any declaration processes. Given the data heterogeneity, sheer volume, and evolving regulatory demands, manual processes have proven inadequate to handle substance identification and classification at scale. This makes harmonizing and enriching these diverse data streams into a unified platform key to delivering comprehensive, consistent substance intelligence.

The AI-powered Siemens Environmental Knowledge Graph (SiEKG) was developed to address this profound, integrating internal and public data sources through machine learning and environmental expertise.

The knowledge graph is a graph-structured knowledge representation system that captures and maps the intricate relationships between substances, their various identifiers, regulatory classifications, and properties. Unlike traditional relational databases, its graph architecture represents entities (e.g., substances, concentrations, suppliers) as nodes, with edges representing the relationship between them, making complex queries straightforward. SiEKG functions as a semantic bridge by mapping disparate real-world vocabulary to standardized taxonomies used in regulatory reporting.



At the same time, it integrates various data sources, including chemical safety data sheets, product data, and purchasing volumes, to trace and statistically estimate substance flows across the entire operation.

SiEKG leverages the combination of the knowledge graph's semantic capabilities, multiple AI models, and explicit uncertainty quantification to automatically reconcile and interpret hazardous substance data from diverse sources. Crucially, the knowledge graph is designed to host quantified uncertainties of predicted imputations. A deep learning algorithm is trained concurrently on existing labeled data to predict likely substance profiles where data is missing.

These AI estimations for missing data are populated back into the knowledge graph along with the quantified uncertainties. The data is then consolidated in the Siemens Data Cloud to compute KPIs to track the environmental performance. This enables systematic estimation of hazardous substance flows across company operations. Environmental experts conduct quality assessments through expert elicitation, and their feedback continuously improves model performance.

A patent filing underscores the novelty of this approach, particularly due to its innovative combination of techniques and use in industrial hazardous substance tracking. While standard industry practices typically rely on simple linear extrapolation from sample data, our multi-layered approach also includes knowledge graphs, multiple AI models, and explicit uncertainty quantification.

The innovation lies not in inventing new algorithms, but in the sophisticated combination and application to a previously unsolved industrial challenge.

In a CSRD environment where data defensibility is a business imperative, SiEKG shifts hazardous substance management from an operational effort to a driver of resilience, auditability, and strategic business advantage. By centralizing data and enriching data features with machine learning estimates into a unified intelligence layer, SiEKG enables us to assess compliance exposures, reduce supply chain uncertainty, and accelerate reporting readiness.

SiEKG processes substance data covering thousands of materials in hours rather than months. More important than speed is scope: The system handles analyses that are impossible manually – tracking trends, identifying patterns, comparing operations, and responding rapidly to regulatory queries with comprehensive data coverage. Building on its existing relations and ontology, it's possible to scale SiEKG to gradually integrate additional environmental data such as recycled material content, material composition, or GHG emissions. It can also connect with other knowledge graphs within the Siemens data ecosystem via aligned ontologies to enable cross-domain insights.

The project demonstrates that machine learning can address fundamental data quality challenges in corporate sustainability reporting: not by mandating clean, perfectly structured data as input, but by working effectively with real-world messiness. Machine learning proves its value by solving problems that rule-based systems alone cannot handle due to inconsistent data, ambiguous rules, and endless exceptions.

From pilots to practice: Transferability and scaling potential

Together, SERA and SiEKG show how AI-powered environmental intelligence can be systematically scaled across our global operations. Their architecture follows a shared design principle: solutions are built to be transferable by default. Once validated in a specific domain, such as hazardous substance tracking or water risk assessment, it's possible to adapt proven AI-powered models in adjacent domains with comparable data characteristics, with careful domain analysis, retraining on target-domain data, and empirical validation before deployment. This enables model re-use and supports the systematic expansion of data intelligence capabilities across business units.

In effect, we are establishing a scalable framework that enables rapid adaptation to new datasets, evolving regulatory demands, and emerging risk domains without the need for fundamental architectural redesign.

Transferable and scalable solutions enable resilience to regulatory change. Adding new disclosures through controlled extensions and targeted retraining instead of complete system rebuilds preserves consistency, avoids one-off solutions, and streamlines complex real-world messiness into audit-ready, reliable data flows.

4 AI as a catalyst for environmentally sustainable impact

The integration of AI and advanced data analytics into environmental protection is reshaping how companies like Siemens manage sustainability – driving not only operational efficiencies, but also enhancing regulatory compliance and disclosure readiness.

Scaling environmental sustainability with data intelligence

The impacts are evident across three main areas: First, advanced analytics enable more precise environmental accounting which empowers more efficient and precise environmental performance management. Second, automation significantly improves data quality and streamlines processes, leading to substantial efficiency gains in both time savings and reduced operational costs. Third, AI-enabled assistance improves easy access to specialized expertise across the organization, while driving adoption through guided processes and synthesized information.

This demonstrates that environmental protection and robust business value are not competing forces, but rather deeply synergistic, with environmental and business benefits not being separate considerations but reflecting different perspectives on the same underlying performance improvements.

Managing AI's footprint

Building on this synergy, AI emerges as a powerful lever to scale sustainability efforts. However, this advancement also necessitates a responsible approach to managing AI's own environmental footprint. Leading responsibly requires actively addressing the impacts of the technologies we deploy, recognizing the inherent links between technological advancement, environmental stewardship, and sustainable economic growth.

„As AI demand grows, we acknowledge the importance of steering it towards Sustainable IT. We analyze its footprint, optimize our own AI solutions and actively engage with our providers to improve our energy and water efficiency."

Helena Babelon, Head of IT Sustainability at Siemens.

Although AI introduces incremental demand to our IT infrastructure, its overall share of the Siemens footprint remains small. Nevertheless, we rely on Siemens' comprehensive renewable energy strategy and energy-efficient infrastructure to balance AI's growing energy needs with our environmental targets. This includes targeted strategies for company-own data centers, as well as partnerships with leading providers of sustainable cloud services, which are essential for deploying our AI workloads. All our internally used AI cloud services run on 100% matched renewable energy.

As technology evolves, particularly in areas like AI and advanced computing, we remain committed to ensuring that innovation consistently aligns with our environmental targets and continues to support both our digital and sustainability transformation.

Looking Ahead: Toward a scalable model of environmental intelligence

Over the next years, environmental data intelligence will mature into a core capability for Siemens – shaping how we plan, operate, and report. The focus will shift from building isolated solutions to embedding them seamlessly into everyday environmental management and reporting processes. AI-supported assessments will increasingly assist our business units by surfacing risks, generating more accurate environmental data at the product and site level, and integrating sustainability considerations directly into planning processes.

Agentic AI systems are expected to play a growing role in this evolution. By orchestrating complex data collection, analysis, and decision-support tasks, these systems can significantly expand our capacity for complex environmental assessments and help deliver more consistent and objective results.

At the same time, this development introduces fundamental questions: How can we maintain management responsibilities when agents act across processes? How can we audit such systems and make their decision paths transparent? How can we continuously monitor and govern their behavior? Addressing these issues will require an evolution of skills, roles, and governance frameworks. Organizations will need to strengthen AI literacy across management and expert functions, establish clear accountability models, and adapt control and assurance mechanisms to ensure that agentic AI operates responsibly, transparently, and in line with regulatory and ethical expectations.

Consequently, environmental intelligence will not only support compliance – it will shape competitive advantage. The ability to generate traceable data at speed, understand environmental implications early, and respond with confidence creates resilience in a landscape of rising expectations and growing complexity.

DISCLAIMER

Siemens AG
Environmental Department
Subject to changes and errors.

The information given in this document only contains general descriptions and/or performance features which may not always specifically reflect those described, or which may undergo modification in the course of further development of the products. The requested performance features are binding only when they are expressly agreed upon in the concluded contract.
All product designations may be trademarks or product names of Siemens AG or other companies whose use by third parties for their own purposes could violate the rights of the owners.

SIEMENS