

Research Data Management

White paper

Tracking institutional research data

How universities can nurture a culture of data sharing





Executive summary

The research landscape is undergoing a period of change in relation to open science, including transparent data sharing, driven by funders, government agencies, publishers and other research stakeholders. To respond to the growing range of drivers and initiatives in this direction, institutions are developing policies, deploying dedicated staff and creating new roles that can cater to research data from its creation all the way through to publication, sharing and preservation. This change, however, does not happen overnight: to strengthen long-term strategies around research data, research stakeholders must nurture a data sharing culture based on openness and trust.

The tracking of research data affiliated with an institution can be a powerful tool in this context: it can signal to researchers that data is seen as a first-class output by decision-makers, as well as underpinning efforts to frame

an institution as a leader in open science. Knowledge of the data produced by affiliated researchers can also help in identifying good practice and enhancing data sharing policies in target disciplines, building positive feedback loops across the institution. Data stewards and data curators can play an important role in enabling the above, thanks to their internal networks and their close understanding of the disciplinary communities they support.

Whether developed in-house or outsourced, data tracking solutions can help institutions monitor progress towards their strategic commitments around research data as well as academic excellence, societal impact, collaboration and partnership. With today's high level of interest in research integrity, transparency and openness, reinforced by policies in the same direction, research data is rapidly rising up the agenda: now is the time for action.



“If researchers don't manage their data responsibly, it might be lost by accident. If you lose research data, your funded project might be delayed, or won't deliver as much value as it should.”

– Maria Cruz,
Senior Policy Advisor, NWO
(Dutch Research Council)

A tale of data loss

In 1946, two professors at the University of Pennsylvania unveiled ENIAC, short for Electronic Numerical Integrator and Computer. ENIAC, as the first general-purpose digital computer, was several orders of magnitude faster than previous electro-mechanical machines, and quickly attracted the attention of scientists, government and business for a number of emerging projects. The transformative nature of early computers was only a hint of the opportunities and applications that would emerge in years to come – something that the research sector benefited from considerably. Developments in computing power and capability have led to novel and increasingly flexible research methodologies and tools, to the point where, today, digital devices are ubiquitous in almost all research activities.

The amount of research data lost or misplaced over time because they weren't properly managed, archived and preserved is so vast it is difficult to estimate.

What the introduction of computing did *not* bring about was a corresponding diffusion of good data management practices. Although the dynamic flexibility afforded via the use of computers, as opposed to traditional pen and paper, is a huge asset, it can be a double-edged sword in a research context. The amount of research data – broadly defined as the results of observations or experimentation that validate research findings – lost or misplaced over time because they weren't properly managed, archived and preserved is so vast it is difficult to estimate.

Over the past few years, there has been increasing interest in tracking research data, whether within an institution (including unpublished and unfunded research data) or after it has been shared in the public domain (e.g., via institutional, generalist or disciplinary repositories). This reflects a growing awareness that data, regardless of one's discipline, is a key research output, complementing the narratives in published articles and enabling others to verify and build on the author's work. However, we are far from the finish line: the journey from data loss towards a full appreciation of research data as a first-class research output is a long and complex one.

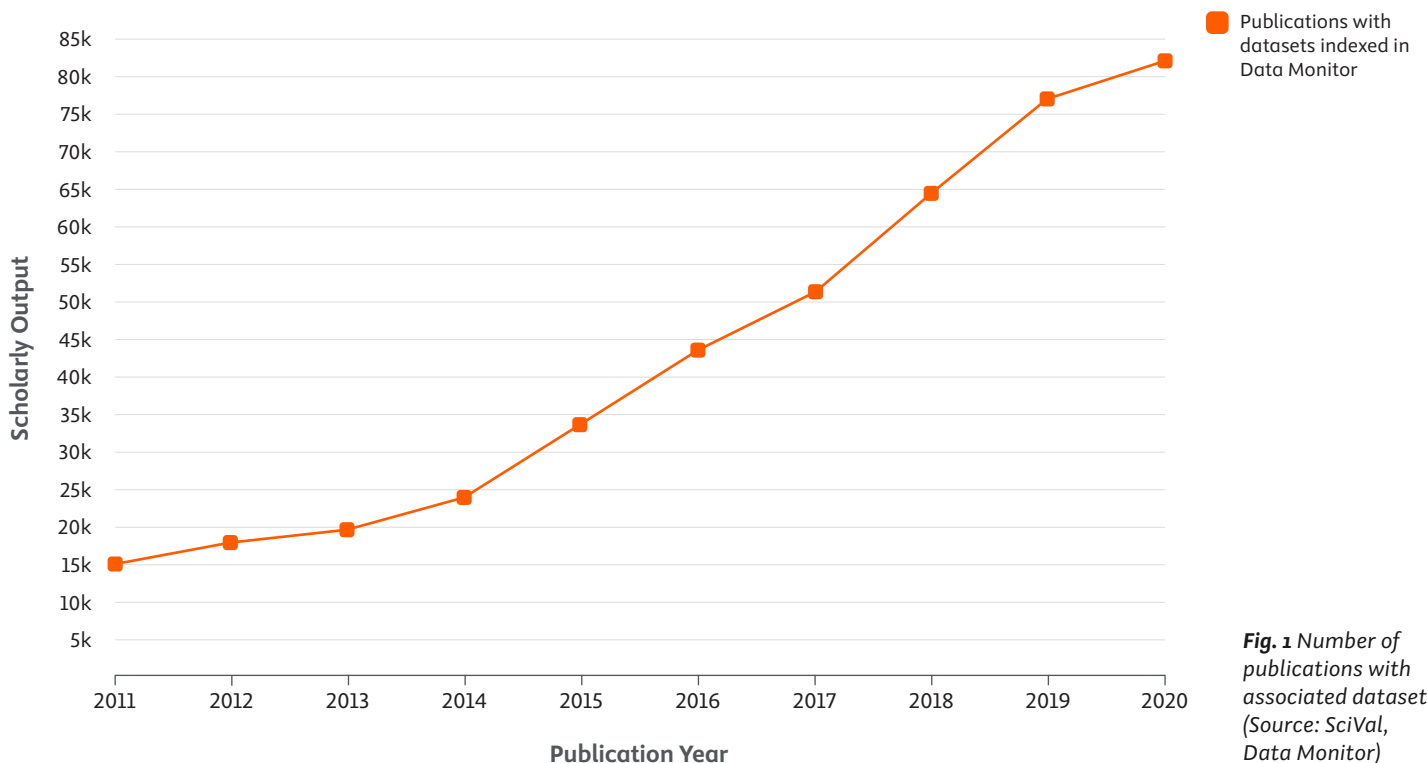


Fig. 1 Number of publications with associated datasets (Source: SciVal, Data Monitor)

Research funders recognize open data sharing as an emerging priority

The ability of researchers to address societal challenges is significantly compromised if there is a lack of trust. In a global context where fake news and malicious players aim to manipulate people's perceptions, research integrity, open sharing and reproducibility are the building blocks on which trust in research depends.

“If you regard research data as a reservoir to respond to crises, these data are the basis to create the knowledge that we need in future. Without data, you couldn't define convincing responses to crises such as the Coronavirus, the climate crisis and the loss of biodiversity.”

– Johannes Fournier, DFG

All stakeholders have a role to play in nurturing data sharing and transparency. For example, government actors and research funders have recognized the important role of open science practices in today's research culture:

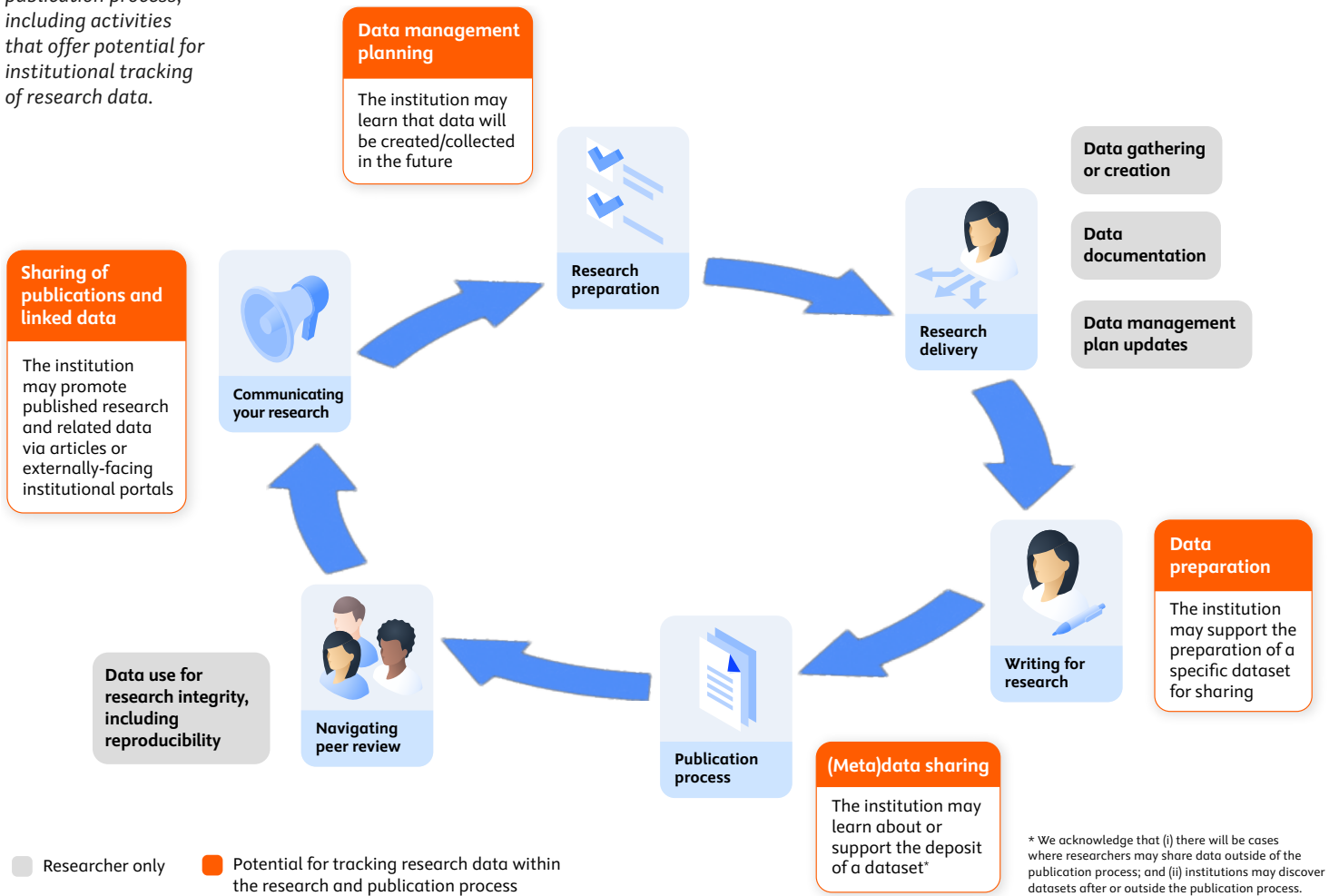
- The [G7 2021 Research Compact](#) speaks of a need “to drive a culture of rapid sharing of knowledge, data, software, code and other research resources”.
- The [US Office of Science and Technology Policy 2022 report on Protecting the Integrity of Government Science](#) emphasizes the need to make “evidence-based decisions guided by the best available science and data, recognizing that scientific and technological information, data and evidence are central to the development and iterative improvement of sound policies and to the equitable delivery of programs across every area” of government.
- The [NIH Data Management and Sharing Policy](#) notes that data sharing “will facilitate the development of treatments and products that improve human health.”
- The [Horizon Europe program](#) includes clusters, partnerships and missions around better interoperability and sharing of data that is findable, accessible, interoperable and re-usable (FAIR).
- Although data sharing is not a cOAlition S principle, the [Plan S implementation guidance](#) “strongly encourages that research data and other research outputs are made as open as possible and as closed as necessary.”

Publishers can support this developing landscape through journal policies, guidelines and tailored workflows. For example, improved data sharing practices can be achieved by interacting with researchers at the right point in the journal publication process, through actions and reminders that encourage and facilitate data referencing or even by directly supporting the deposit of datasets.

From an institutional perspective, it is complex to reconcile these external policy requirements where the responsibility lies with individuals with the support and training that libraries and other professional services can provide. For example, there are several stages in the research and publication process where institutions could learn whether researchers will collect or produce data that they can share. Institutional support for data management planning when researchers are applying for research funding is an important early-stage touch point. In addition, continuing engagement throughout the research lifecycle is needed to ensure that the reality of project delivery fully reflects the planned approach to research data management. However, such an approach relies on researchers themselves getting in touch with professional services for support or advice. In all other cases, it is typically difficult for an institution to become aware of and track research data produced by affiliated individuals, particularly where this is after, or not within the context of, the publication process.



Fig. 2 Research and publication process, including activities that offer potential for institutional tracking of research data.



Tracking research data

Tracking research data refers to the identification of the data output of a given institution, to build or improve a data catalogue or data registry. This can be achieved in various ways, ranging from purely manual (and, therefore, labor-intensive) efforts, to the development of dedicated data tracking solutions in house, to the procurement of third party data tracking solutions. Tracking efforts would typically begin by identifying a corpus of relevant

research data, for which metadata records would need to be gathered, cleaned and enriched before they can be used. For example, metadata records from different sources need to be deduplicated and improved, with a key barrier often being the low quality of institutional affiliations at the time of data deposit. If desired, one may also link datasets to research funders and specific grants, potentially building on digital identifiers where these are available.



Institutional strategies must nurture a data sharing culture

Institutions are moving from aspirational approaches that engage with the high-level principles of data management to embedding the practice of data management and governance in strategies for excellence in research. It is increasingly common for research performing organizations to reflect research data as part of their strategic goals, particularly in terms of their ability to publish trusted work of the highest quality and ambition. This has significant repercussions for broader research practice and organizational culture, and there has been growing recognition that research data sharing and tracking are not solely technical endeavors. To enable these practices, the right mix of people, internal knowledge and resource sharing channels, IT and software provision, funding and, importantly, genuine commitment to research integrity at every level, are key.

“Over the years, I have talked to many institutional administrators who led impactful initiatives to advance RDM but shared a challenge: how to keep track of their research data. This was described as a key step in understanding where their institution was in its RDM journey, and limited knowledge in this area was seen as a barrier to assessing progress.”

– Lorenzo Feri, Elsevier

Establishing effective research data management and sharing as a core element of a positive research culture is a journey that all stakeholders must undertake together. In this context, the tracking of institutional research data can help research managers and leaders develop reports and metrics that apply to different faculties or departments, and usefully ground a discussion that would otherwise be hypothetical.

In addition, the communication of scientific findings to the public, policymakers, industry and any other interested parties is becoming more central in institutional strategies. For example, the ability to highlight and share data created by affiliated researchers (e.g., via dedicated institutional portals) can underpin the creation of policy impact, new collaborations and creative and unexpected (re)uses of research data.

The right mix of people, internal knowledge and resource sharing channels, IT and software provision, funding and, importantly, genuine commitment to research integrity at every level, are key.



“Among other things, knowledge of data created by researchers is invaluable in identifying research gaps to drive program priorities. Data sharing is a priority at NIH and our goal is to shift the culture to make data sharing the norm. Planning to adopt best practices early in the research lifecycle will bring recognition and a return on the investment not only to the PI and institution, but also to the funders of the work.”

– Ishwar Chandramouliswaran,
Program Director, NIH Office
of Data Science Strategy,
National Institutes of Health (NIH)



By tracking institutional data assets, teachable moments as well as impactful case studies can be more easily surfaced, with a positive impact on both internally and externally facing communications.

Proactive support on research data sharing is crucial

Professional support offered by data stewards or data curators offers a way of embedding proactive data expertise as an element of an effective research culture. This can help to fill knowledge or specialist expertise gaps that researchers may have in this relatively new and continually evolving space.

Strategic investment in roles such as data stewards and data curators is growing in popularity and stems from the need to address questions of rights, roles and responsibilities in relation to research data management. Research performing organizations are also recognizing overlaps between data management and privacy, data protection and intellectual

property considerations. For example, information, guidance and requirements relevant to research data management may be included in policies relating to research ethics, governance, integrity, data security and information governance.

A core challenge for institutions, libraries and administrators is to find impactful examples and case studies that enable them to demonstrate the importance of effective research data management and sharing. By tracking institutional data assets, teachable moments as well as impactful case studies can be more easily surfaced, with a positive impact on both internally and externally facing communications.

Individual practices strongly affect an institution's ability to track data assets

Individual researcher attitudes towards open research data sharing vary between strong support – a cohort of people who are open research practitioners and want to share – to researchers who consider data as their own asset that is not suitable for broader dissemination, to others who are keen to share in principle but are concerned about appropriately protecting sensitive data from unscrupulous actors. Disciplines where one should, indeed, think twice before openly sharing data abound: for example, when it comes to research on endangered species or modern slavery, [open sharing may give rise to significant risks](#). Either way, research data in any domain should be proactively managed and, where appropriate, shared, including the use of appropriate technological solutions in line with ethics requirements (e.g., access controls, trusted research environments). Importantly, institutions should be aware of the reasons for decisions on open vs controlled-access sharing – not least because of the potential risks to the organization if the management of these data were to go wrong.

“We have data stewards, our disciplinary data management experts, who focus on requests from their own faculty. They are very familiar with the research context in that faculty, the culture and behavior of researchers and the practical challenges.”

– Yan Wang, TU Delft

More fundamental disciplinary differences also affect individual attitudes towards data sharing. A common challenge is the fact that, in the Arts and Humanities, researchers may not think of the artefacts of their research as “data”, potentially leading to limited sharing and opportunities for tracking. Data in the Arts and Humanities may take many forms, including digitized archives, diaries, pictures, or customized databases built for a specific research project; a tailored language may be more appropriate for engagement around sharing and preserving some of these materials, and for ensuring that the institution is aware of and is sharing as much of its research data output as possible.

Notably, disciplinary differences also affect repository choice, including institutional, generalist and disciplinary repositories based within one's country or managed as international resources. Research on a sample of institutions found that [up to 90% of research data](#) may be deposited outside of institutional repositories, across thousands of databases collectively adopted by research communities or recommended by research funders. The use of disciplinary repositories typically leads to good practice (e.g., the use of discipline-specific metadata) but also makes it more difficult for an institution to become aware of and track data shared by affiliated authors.



“We would like to be able to pull in all research data with an Oxford lead or collaborator into some sort of registry, ideally via our institutional repository, so that we know what Oxford related data exists and where that data resides, even if we don't necessarily own it.”

– David Tomkins,
Research Data Curation
Specialist, Bodleian Libraries,
University of Oxford

Research on a sample of institutions found that up to 90% of research data may be deposited outside of institutional repositories, across thousands of databases collectively adopted by research communities or recommended by research funders.



Institutions need to balance people and technical infrastructure for tracking research data

In a research culture conventionally focused on publications as the definitive output of research, fewer researchers will be likely to deposit data openly in an online repository and will prefer instead to use personal or institutional data storage locations. Furthermore, time-pressed researchers may want to focus on the next bid, grant or project, rather than revisiting data management requirements from a previous project.

The fact that data sharing lies at the intersection of good research practice, research integrity, disciplinary norms, external requirements and individual belief in open science often makes institutional intervention difficult. For example, researcher decisions about if and where they deposit their data are frequently invisible to their organizations, unless they deposit their research data (or, at least, metadata) in the local institutional repository.

“There can be resistance from researchers to sharing data, but researchers, managers and leaders all see, in principle, the importance of research data.”

– Dongmin Seo, Korea Institute of Science and Technology Information

Given the high variation in behaviors within a single organization, institutions may struggle to prioritize resources such as training and support, but also IT spend, across their research portfolio as well as their departments and faculties. Improved knowledge of the institutional research data landscape could certainly help identify which departments or faculties should be targeted or need dedicated intervention, with a potential impact on the achievement of strategic priorities. In addition, without reliable evidence or reporting on research data management and sharing practices, the above concerns are easy to deprioritize, as they can't be quantified and addressed via targeted interventions. In simple terms, if institutions struggle to gauge the state of their current research data management and sharing landscape, they are not in a position to respond fully to the needs and desires of the academic community.



“For us, research data management is about research integrity. We recognize that accurate and retrievable research data are an essential component of any research project and necessary to verify and defend when required, the process and outcomes of research. This is why we highly encourage our researchers and post graduate students to submit their research data to us.”

– Jesse Xiao,
Head of Scholarly
Communication and Research
Services & Medical Librarian,
Hong Kong University

Exploring use cases for research data tracking

Before starting to track research data, it is important to consider specific use cases that this can help institutions address. This is an essential step in securing buy in and funding within an institution: regardless of what data tracking solution is chosen, a business case is likely to be needed to justify the investment in people and technology.

Institutions may use the outputs of research data tracking for a broad range of purposes, based on the granularity of the information (i.e., metadata) collected. For example, use cases may include:



The identification of pockets of data management and sharing excellence



The targeted enhancement of data sharing practices with the support of dedicated personnel (e.g., data stewards or data curators)



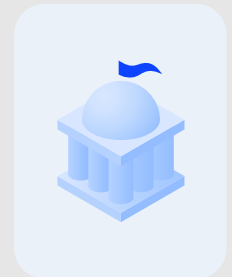
Compliance checks against institutional or funder data management policies



The preparation of reports for senior leaders around strategic KPIs on research data



The publication of institutionally affiliated research data via a public portal



The promotion of the institution as a global leader in the open science landscape

Additionally, the above use cases all work in unison to help institutions enhance the standing of research data as a first class research output, contributing to furthering discussions around research culture and transparency in the context of research integrity and good research practice.

The tracking of research data is underpinned by emerging practices

The scholarly communication infrastructure that can inform the tracking of research data is highly distributed. Data may be found in a complex network of institutional, generalist and disciplinary repositories, but also in other forms such as on individual websites that can only be discovered via tailored approaches due to limited or lack of metadata or machine-readable information. The University of Groningen has described the use of software solutions to identify institutionally affiliated research data as “[systematically collecting needles from the haystack](#).” In their case, the use of [Data Monitor](#) helped cut the time spent validating dataset metadata from 20–30 to 5–10 minutes and allowed the University to grow its list of validated datasets from about 600 per year (gathered manually) to almost 4,000 per year (gathered via software and then validated).

Institutions, researchers, publishers and funders need to continue working together to foster improvements in individual behaviors, which will eventually lead to an enhanced ability to track research data. The above-mentioned FAIR principles offer a helpful focal point that all stakeholders can consider, from the individual researcher all the way through to international digital infrastructures. For example, those involved in data production should ensure that they deposit their data in an appropriate repository, which is in turn expected to accept (and, in some cases, require) complete, accurate and meaningful metadata and controlled vocabularies, as well as to create and manage a persistent identifier.

The use of persistent identifiers – whether digital object identifiers (DOIs), permalinks or others – is key in enabling institutions to identify data shared by their researchers and improving findability more broadly.

“In many cases, information about affiliation is missing from metadata. And, when it is there, the persistent identifier of the organization is only rarely available. Persistent identifiers could help overcome some difficulties with research data tracking, but they’re not widespread enough today.”

– Paolo Manghi, OpenAIRE

These don’t only include DOIs or other persistent identifiers for datasets, but also Open Researcher and Contributor ID (ORCID) and Research Organization Registry IDs (ROR) as a minimum. Other forms of identifiers are also emerging, for example Research Activity Identifiers (RAIDs) for projects and Crossref grant DOIs.

Furthermore, organizations responsible for repository management can ensure data is curated and preserved, that metadata is surfaced appropriately and that the repository is accessible and optimized for discoverability.

“We gather all successful research data sharing stories on the NYCU Dataverse knowledge website, aiming to increase reuse and enhance international cooperation.”

– Ming-Jiu Hwang, National Yang-Ming Chiao Tung University

The other side of the coin is people involved in data consumption, including authors themselves. Today, data users are more frequently ensuring that they accurately cite and link to third-party data, and may incorporate linking to data assets in narrative CVs.

However, data citation practices are far from mature. For example, publications may include data citations as part of the full text, in dedicated data availability statements or not at all. This kind of fragmentation is partly responsible for the difficulties in tracking research data and sometimes leads to the need to disambiguate, deduplicate or further process the information that can be harvested from the scholarly infrastructure.



“We aim to get to a point where we can do more with our research data and certainly track it. We are very aware that there is more that could be done but getting to that tipping point where things are all joined up does take time.”

– Isobel Stark,
Head of Research Data and Intellectual Property,
University of Southampton



Institutions, researchers, publishers and funders need to continue working together to foster improvements in individual behaviors, which will eventually lead to an enhanced ability to track research data.

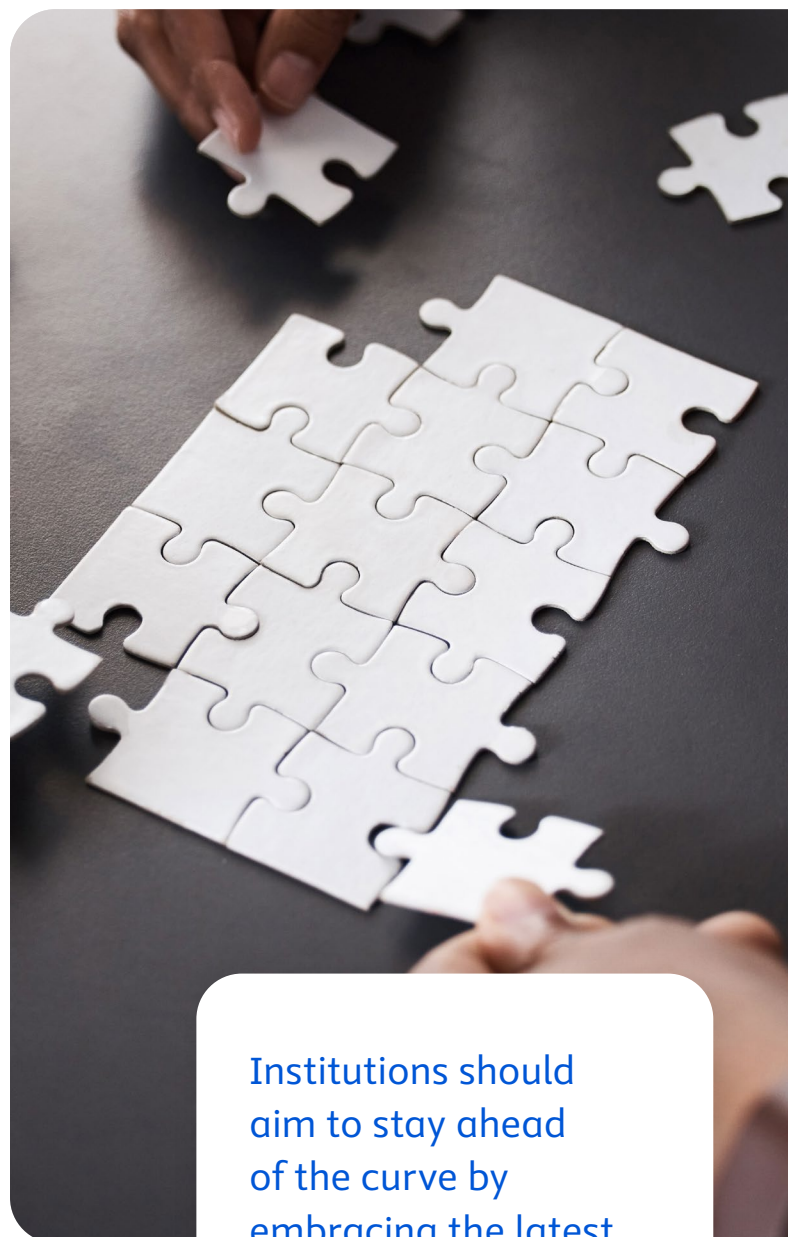
The rationale for action

Effective data management and sharing are a cornerstone of excellence in research culture. These practices, as part of the broader research integrity agenda, help institutions enhance trust in their published findings as well as more incisively showcase impact and public benefit from research.

As a key stakeholder in the research community, institutions are increasingly interested in tracking the data collected or created by affiliated researchers. This knowledge can help inform support and training provision as well as the prioritization of IT spend based on an enhanced understanding of local communities of practice. Risk and reputation management are enabled by the tracking of research data, too, allowing institutions to simultaneously strengthen their provision and minimize threats. Finally, improved knowledge of institutional research data sharing can be a significant asset in showcasing societal impact and in securing partnerships and collaborations.

While the research data sharing landscape is still somewhat in its infancy relative to traditional publication outputs such as research articles, concerted efforts are being made to improve the extent to which research data can be monitored and tracked. The increased use of persistent identifiers and data citation practices, the consideration of data in peer reviews and the requirement for data availability statements by research funders and journals are just a few examples of cross-stakeholder initiatives that support the growing prominence of research data.

Institutions should aim to stay ahead of the curve by embracing the latest developments in the research data landscape and by joining a growing number of peers in publicly showcasing their commitment to transparency, integrity and excellence in research.



Institutions should aim to stay ahead of the curve by embracing the latest developments in the research data landscape.

A 10-step plan to begin tracking your institutional research data

The discussion presented so far is high-level and covers a broad landscape of stakeholders, behaviors and developments. Building on these insights and the views of the experts consulted as part of this research, we provide some practical next steps that institutions can pursue. The following plan is a series of recommended steps that can be adapted by institutions looking to strengthen their approaches to (open) research data management and sharing and move towards data tracking.



1 Review your current support for open science practices.

As a starting point, consider where your institution stands with regard to open research, including policies, practices and behaviors. This will help frame any discussions around (open) data management and sharing and its subsequent tracking.

2 Develop a mission statement around managing and tracking research data.

Think about the level of commitment you wish to make towards data management and tracking. For maximum impact, you may include a commitment to research data as part of the institutional research strategy, but strengthening an existing research data management or governance policy may also be effective depending on your local context.

3 Establish roles and responsibilities, including a lead on data management and tracking.

To drive progress and build momentum, make sure a working group or similar structure is established, including a lead individual, clear roles and responsibilities and appropriate decision-making structures. Given the impact of data tracking on the institution as a whole, it is essential to engage a range of functions, for example IT services, research services, the institutional library, the data protection office and data ethics and governance.

4 Review budget and resourcing available to achieve your strategic ambition.

Assess the budget and staff time available within your institution and begin to form a view as to whether it might be desirable to outsource a portion or the entirety of your data tracking efforts. In some circumstances, you may need to prepare a business case for the creation or procurement of a data tracking solution. The contents of this white paper are likely to help you shape up your business case and secure buy-in from senior leaders.

Cont.



5 Review the data tracking solutions available, including in-house delivery and outsourcing.

Data tracking involves a breadth of activities, starting from coding and moving on to curation, deduplication, linking and cleaning. In addition, it is based on somewhat complex technical infrastructure, such as a data warehousing solution. To move forward, institutions should review the technical solutions already available in the landscape and compare them against what might be achievable in-house.

6 Consider the extent of integration needed with existing institutional software.

Today, institutions are likely to use a Research Information Management System (RIMS) or Current Research Information System (CRIS), including to track open access publishing. It is recommended that the extent of integration between the existing RIMS or CRIS and the planned data tracking system be considered.

7 Shortlist and trial a small set of options and their ability to meet your objectives.

Schedule demos and hands-on sessions with existing tools and/or run some tests to see what your institutional capability may be best placed to deliver. Critically compare these options based on how well they meet the objectives set by your institution. In addition, explore unexpected functionality or features that you were not aware of but might add value for your institutional and external stakeholders.

8 Choose and implement your data and tracking solution.

Upon consultation with the group driving your institutional data tracking efforts, choose a solution (whether in-house or outsourced) and establish clear operational responsibilities for its launch as well as a target timeline. If delivering data tracking internally, you may need to liaise with other institutional functions to carry out recruitment and the procurement of new technical solutions. If outsourcing your data tracking efforts, develop or sign a service agreement and put in place appropriate schedules for payment and timelines for renewal.

9 Train and educate your staff, including academics as well as professional services.

Develop communications and engagement activities to highlight the added value provided by data tracking to staff across the institution. At the same time, ensure that a clear and unambiguous set of individuals are seen as the service owners, so that queries can be directed to them. Ideally, staff support on data tracking should be mirrored by support throughout the research process, e.g., via data stewards or data curators.

10 Audit the chosen data tracking approach and verify its performance and reliability.

At the end of the implementation process, ensure that your chosen solution performs as needed: for example, are results credible and trusted by their target audience? In addition, consider whether it is helping achieve your institutional strategic targets. Keep performance under review, as well as the landscape of solutions available, as scholarly metadata is evolving rapidly and new and improved data sources emerge frequently.



Acknowledgements

This white paper was built on extensive desk research as well as via consultation with experts based in organizations across the world. The authors thank the following individuals who kindly contributed to our consultation:

Name	Affiliation and role	Country
Alan Slevin	University of Strathclyde Glasgow, Scholarly Publications & Research Data Manager	UK
David Robertson	University of Edinburgh, Professor of Applied Logic, Informatics; Co-Director, Centre for Medical Informatics	UK
David Tomkins	University of Oxford, Research Data Curation Specialist	UK
Dongmin Seo	Korea Institute of Science and Technology Information (KISTI), Director of Research Data Sharing Center (Division of National Science & Technology Data)	South Korea
Hanna Shmagun	Korea Institute of Science and Technology Information (KISTI), Researcher at NTIS Centre (Division of National Science & Technology Data)	South Korea
Ishwar Chandramouliswaran	National Institutes of Health (NIH), Program Director, NIH Office of Data Science	USA
Isobel Stark	University of Southampton, Head of Research Data and Intellectual Property	UK
Jaana Pinnick	British Geological Society; National Geoscience Data Centre (NGDC), Grants Manager and Research data and digital preservation manager	UK
Jesse Xiao	University of Hong Kong Libraries, Head of Scholarly Communication and Research Services & Medical Librarian	Hong Kong
Johannes Fournier	DFG, Programme Director, Scientific Library Services and Information Systems; Expert Committee for National Research Data Infrastructure	Germany
Judith Carr	University of Liverpool, Research Data Manager	UK
Laurian Williamson	University of Leicester, Associate Director and Head of Research Services	UK
Lorenzo Feri	Product Director, Elsevier	Netherlands
Maria Cruz	NWO, Netherlands, Senior Policy Advisor	Netherlands
Martin Wolf	University of Liverpool, Head of Open Research	UK
Ming-Jiu Hwang	National Yang-Ming Chiao Tung University, Library Director	Taiwan
Paolo Manghi	OpenAIRE, Chief Technical Officer	Italy
Radek Pajor	University of Leicester, Research Services Consultant (Research Data Management)	UK
Sarah Stewart	University of Oxford, Head of Research Data Management Provision	UK
Yan Wang	TU Delft, Data Stewardship Coordinator	Netherlands

References

1. Freiberger, Paul A. and Swaine, Michael R. "ENIAC". Encyclopaedia Britannica, 12 Jan. 2023, <https://www.britannica.com/technology/ENIAC>. Accessed 23 February 2023.
2. Cabinet Office. "G7 2021 Research Compact", 12 Jul. 2021. <https://www.gov.uk/government/publications/g7-2021-research-compact/g7-2021-research-compact>. Accessed 23 February 2023.
3. Scientific Integrity Fast-Track Action Committee. "Protecting the Integrity of Government Science". Jan. 2022, https://www.whitehouse.gov/wp-content/uploads/2022/01/01-22-Protecting_the_Integrity_of_Government_Science.pdf. Accessed 23 February 2023.
4. National Institutes of Health. "2023 NIH Data Management and Sharing Policy". 25 Jan. 2023, <https://oir.nih.gov/sourcebook/intramural-program-oversight/intramural-data-sharing/2023-nih-data-management-sharing-policy>. Accessed 23 February 2023.
5. European Commission, Directorate-General for Research and Innovation. "Horizon Europe, open science: early knowledge and data sharing, and open collaboration". Publications Office of the European Union, Jun. 2021, <https://data.europa.eu/doi/10.2777/18252>. Accessed 23 February 2023.
6. cOAlition S. "Part II: Guidance on the Implementation of Plan S". n.d., <https://www.coalition-s.org/guidance-on-the-implementation-of-plan-s/>. Accessed 23 February 2023.
7. The University of Manchester & University of Groningen (2023). Needle in a Haystack: Where is my Institution's Data? Monitor & report research data. https://elsevier.widen.net/s/nvsgqxggp6/rdm-webinar-2_presentation_lib_ri_web
8. Zudilova-Seinstra, E. (2020). Analysis of research data for 11 Institutions - Data Monitor. [Data set]. Mendeley. <https://doi.org/10.17632/K5P45Z33KB.3>
9. Open data offer risks and rewards for conservation. (2018). Nature, 559(7715), 444–444. <https://doi.org/10.1038/d41586-018-05800-y>

Data Monitor

Data Monitor puts institutions in the driving seat of their Research Data Management journey by providing visibility on their institution's entire research data output.

Scope

This 2023 white paper, based on an investigation by Research Consulting for Elsevier, explores the rationale for institutional data tracking in the context of open science. Based on a consultation with over 20 international stakeholders, the white paper discusses the policy and practice landscape and provides guidance for institutions looking to strengthen their approaches to (open) research data management and move towards data tracking.

For enquiries about the white paper, please contact:

Andrea Chiarelli

Principal Consultant

RESEARCH CONSULTING

The Ingenuity Centre, University of Nottingham Innovation Park,
Triumph Road, Nottingham, NG7 2TU, United Kingdom

andrea.chiarelli@research-consulting.com

+44 (0)115 896 7567

www.research-consulting.com

For enquiries about Elsevier's data tracking solutions, please contact:

Manik Sethi

Senior Manager, Portfolio Marketing

ELSEVIER

Radarweg 29, Amsterdam, 1043NX, The Netherlands

m.sethi.1@elsevier.com

www.elsevier.com/solutions/data-monitor

Elsevier offices

ASIA AND AUSTRALIA

Tel: + 65 6349 0222

JAPAN

Tel: + 81 3 5561 5034

KOREA AND TAIWAN

Tel: +82 2 6714 3000

EUROPE, MIDDLE EAST
AND AFRICA

Tel: +31 20 485 3767

NORTH AMERICA,

CENTRAL AMERICA

AND CANADA

Tel: +1 888 615 4500

SOUTH AMERICA

Tel: +55 21 3970 9300

DISCLAIMER

The views expressed in this publication are those of the authors. They do not purport to reflect the opinions or views of Elsevier. The investigation was sponsored by Elsevier, who assisted with access to clients and resources to complete the research.