


# IDCC27 Accepted Lightning Talk Abstracts

**A Natural Evolution Toward FAIR Principles: Insights from a Small Quebec Research Institute in a Global Research Ecosystem**

The size of an organization and the resources made available to it are not always aligned with its ambitious institutional objectives. A small institution may nonetheless aspire to make its scientific output findable, accessible, interoperable, and reusable—especially when its research findings directly address specific, real-world needs. Such is the case for the Institut de recherche Robert-Sauvé en santé et en sécurité du travail (IRSST), a French-language research institute specializing in occupational health and safety and based in Montreal, Quebec, Canada.

To meet both institutional objectives and external users' research access needs, the IRSST documentation centre team (composed of 2 librarians and 2 library technicians) launched an institutional repository, PhareSST, in the summer of 2024. This new and unique platform brings together the Institute's entire body of scientific output. Its implementation was a first step, but simply populating the repository was not enough. We needed to maximize its impact by enhancing metadata as effectively as possible.

This presentation explores the evolving discoverability strategy of a small organization, made possible despite limited resources (number of employees, initial lack of expertise in data curation, etc.). The steps we took, feeling our way through the unknown, gradually took shape, as one action led to another, little by little. At every stage, library technicians (as learners) found themselves at the heart of a bold yet necessary data curation effort.

Through reading, self-directed learning, exploring new approaches, and collaboration, we progressively and naturally aligned our practices with FAIR principles. This presentation demonstrates that human input—with the help of tools, not the other way around—is at the heart of operationalizing the FAIR principles. It also outlines the various initiatives implemented at the IRSST to illustrate that people-centered data curation is still possible and necessary, with the aim of inspiring similar institutions to democratize scientific knowledge.



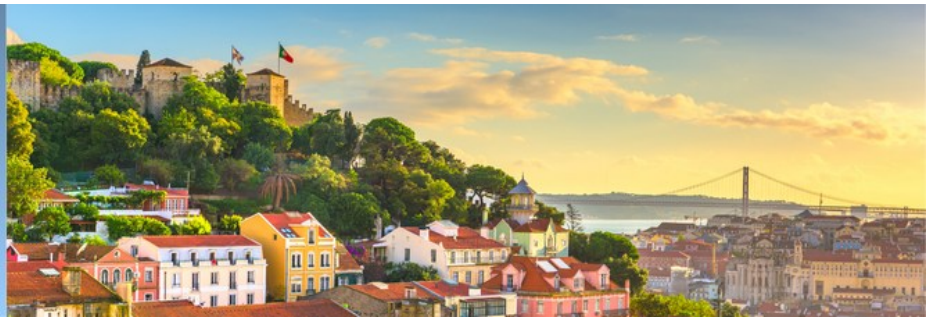

**Advocating FAIR Principles: Training and Informational Support in Research Data Management**

FAIR data and digital objects that cross disciplines are required to solve complex challenges faced by governments and society. These challenges include environmental pressures, like climate change, pollution, and biodiversity loss. Scientists need to be equipped with the skills and knowledge to produce FAIR data resources and digital objects. At the UK Centre for Ecology & Hydrology (UKCEH), a team of Data Stewards provides the informational support to upskill the scientists. We provide support at three levels: a bespoke training programme, structured support to teams and high-impact projects, and ad-hoc support to projects and individuals. This way we advocate for practices that align with the FAIR principles. Firstly, our training consists of different modules relevant to different stages of the data lifecycle. We use a mix of lecture-style teaching and demonstrations to explain important concepts, and interactive exercises to engage learners, embed the learning, and foster discussion. Exercises range from quick polls that gauge knowledge to facilitated break-out discussions which encourage peer-to-peer learning. We centre such exercises around realistic scenarios. Learners are encouraged to propose and discuss potential actions to ensure data are managed well in the presented situation. Secondly, the Data Stewardship Team provide team- and project-level support, as it allows to address issues specific to these teams and projects. Lastly, ad-hoc support to projects and individuals offers another avenue, for more in-depth discussion and personalised support. This three-tiered approach to informational support has expanded our reach and impact within the organisation and has enabled scientists to understand the value of and produce FAIR data and digital objects. This progress is crucial in addressing the environmental challenges faced by humanity today.

**Beyond a Precedent Library: Designing an Indigenous Agreements Database for Context, Authority and Decision Support**

Indigenous land and cultural rights agreement information is increasingly held in digital systems, but too often those systems are designed around storage, retrieval and precedent rather than the people who must use the information to make decisions. This presentation shares early design lessons from development of an indigenous agreements database with a focus on how digital curation can support Indigenous and Traditional Owner Corporations and people, researchers and practitioners without flattening legal, cultural or governance context.

The project sits within a staged twelve-month work program involving data remediation, content accuracy review, copyright and Indigenous Cultural and Intellectual Property assessment, agreement categorisation, user notices, governance protocols, interface design, testing and handover. It highlights a practical challenge for FAIR digital objects as



agreement records may be findable and accessible, but they are not necessarily reusable or interoperable in a safe, culturally legitimate or decision-supportive way unless the surrounding authority, provenance, permissions, restrictions are curated with equal care for context and nuances.

The presentation argues that an indigenous agreements database should be framed as a decision-support and capability-building platform, not a precedent library, compensation benchmark, clause bank or template generator. It should help users understand options, identify risks, ask better questions and recognise when specialist legal, tenure, commercial, cultural or governance advice is required.

Drawing on user-centred design observations from a Traditional Owner Corporation perspective, the talk identifies practical design features including guided search, plain-language layers, contextual warnings, access controls, audit trails, take-down processes, ICIP notices and ongoing Indigenous Data Governance. The core lesson is that centring people in digital curation requires more than making records available. It requires designing systems that protect cultural authority, preserve context, highlight nuances, and support better decisions.

**Beyond FAIR: Integrating Trust, Accountability, and Human Stewardship Through FAIR-A+**

The FAIR Principles have transformed the management and sharing of research data, providing a powerful framework for discovery, reuse, and interoperability. As FAIR increasingly expands beyond research environments into government, cultural heritage, archives, information governance, and artificial intelligence ecosystems, an important question emerges: are FAIR principles alone sufficient to support trustworthy digital objects in all contexts?

This presentation argues that while FAIR successfully addresses discoverability and reuse, it does not fully address questions of trustworthiness, accountability, provenance, and evidentiary value that are critical when digital objects function not simply as data, but as records documenting actions, decisions, rights, and obligations.

Drawing on records management principles established in ISO 15489-1 and perspectives from archives and information governance, the paper proposes an extension to FAIR: FAIR-A+, where the additional "A" represents Authoritative. Authoritativeness encompasses authenticity, reliability, integrity, and enduring usability - characteristics long recognized as necessary for information to be trusted and used as evidence.




The presentation positions FAIR-A+ as a human-centred complement to FAIR Digital Objects. It highlights how stewardship practices that support provenance, context, accountability, and documentation ultimately serve people: researchers seeking reproducibility, citizens seeking transparency, Indigenous communities asserting control over knowledge and data, organizations making critical decisions, and AI systems reliant on trustworthy information.

The paper also explores practical mechanisms - including dataset documentation, provenance metadata, records management controls, and digital curation practices - that can help bridge the worlds of data management and information governance.

Rather than moving beyond FAIR, FAIR-A+ seeks to strengthen it by connecting FAIR Digital Objects to the social, organizational, and ethical dimensions of trust. In doing so, it reframes digital curation not merely as a technical challenge, but as a stewardship responsibility centred on the people who create, depend upon, and are represented by digital objects.

**Beyond the Technical: Service Readiness Levels as a Tool to Assess Service Maturity on Multiple Dimensions**

To assess the maturity of services in digital curation and research data infrastructure, readiness scales such as the Technology Readiness Levels (TRL) are commonly used, for example in the context of the European Open Science Cloud. Although TRL was originally developed for hardware and engineering contexts, it is now widely applied also to services that are socio-technical in nature. This paper argues that such services warrant a more complex approach than a purely technical lens allows.

This talk reports on the development of a Service Readiness Level (SRL) framework as a response to this challenge in the context of a national research data infrastructure programme The SRL aims to re-centre maturity assessment on the human and organisational contributions.

In developing the framework, we reviewed existing readiness and maturity models, from adaptations of TRL for public-sector and non-technical contexts to innovation readiness scales and maturity models specific to research data management. None captured service maturity on the multiple dimensions that it has: existing models tend to be too narrowly technical, oriented toward commercialisation, or focused on institutional capacity rather than the individual service. The SRL framework combines these approaches while adapting it to a research infrastructure context.




It retains the nine-level structure familiar from TRLs but reimagines each level around the maturity of a service as a whole, assessed through qualitative evidence criteria. These reach well beyond technical readiness into user need, adoption, operations, and governance. SRL can be employed through a self-assessment instrument: development teams assess their own maturity, supported by evidence, rather than being subjected to external technical audit.

We share early lessons from validating the framework with real services and invite discussion on how readiness models can recognise, rather than marginalise, the human labour of digital curation.

**Bridging the FAIR Last Mile: Design Principle for a Hybrid AI-Steward Clinic at TU/e**
TU/e has invested heavily in research data management infrastructure, including the Research Cockpit platform for Data Management Plans (DMPs) and steward support. Yet a May 2026 review found only 22% of 1,228 reviewed DMPs intended FAIR-compliant data deposit, revealing a "last mile" gap: researchers understand FAIR principles but struggle to apply them, while stewards can point toward resources but often lack deep, case-based expertise.

To address this, we designed a hybrid FAIR Data Clinic combining AI-assisted assessment with human counselling, structured in five steps: (1) researchers initiate contact, preserving their autonomy; (2) stewards review the DMP for context and catch misalignments; (3) an AI tool subject to appropriate privacy and governance review (and configured using DANS's FAIR-Aware criteria) generates a preliminary FAIR gap analysis as steward preparation—not a final verdict; (4) steward and researcher meet one-on-one to discuss findings and negotiate concrete improvements around metadata, documentation, licensing, and repository choice; (5) a follow-up four/six weeks later checks implementation, with F-UJI scoring the deposited dataset where possible.

The design deliberately keeps AI in a preparatory, training role rather than facing the researcher directly, reflecting a belief that durable FAIR improvement comes from human dialogue attentive to disciplinary context. The clinic also functions as professional development for stewards, building their expertise through repeated practice.

We note this approach trades scale for depth: each case demands two-three hours of steward time, raising questions about institutional resourcing, and the opt-in model may favour already-motivated researchers over those most in need of support.




By the February 2027 conference, the clinic will have run for several months. The presentation will share either quantitative outcomes (session counts, FAIR gaps, F-UJI scores, deposit rates) or, if uptake is low, a reflective analysis of the design process and what it reveals about institutional readiness for FAIR practice.

**Building a European Agricultural Data Steward Network for Trusted Research Data Management Support**

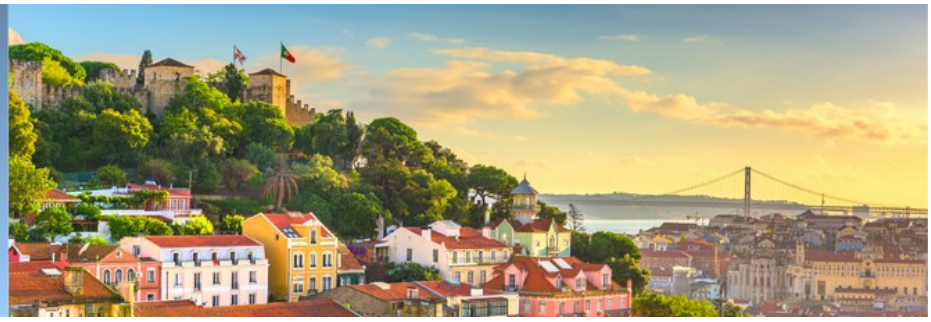
Agricultural research relies on increasingly diverse and interconnected data, from environmental observations and field experiments to geospatial, sensor and socio-economic data. Researchers must integrate these heterogeneous resources while complying with FAIR principles, disciplinary standards and evolving data policies. As a result, research data management support requires specialised expertise that often extends beyond the capacity of local data stewardship services.

To address this challenge, we present a European network of agricultural data stewards and national contact points that connects distributed expertise across institutional and national boundaries. The network currently brings together partners from Germany, Sweden, the Netherlands, Austria, Switzerland and Ireland, with further expansion planned. Through a shared collaboration platform, user requests can—with the researcher's consent—be forwarded anonymously or directly to the most suitable expert. Responses are returned through the local contact point, preserving trusted local relationships while extending expertise across Europe.

Rather than presenting another technical solution, we argue that this network demonstrates the value of human-centred stewardship. It enables knowledge exchange among data stewards, increases the visibility of specialised services and repositories, and strengthens a sustainable community of practice. These human connections provide trusted, quality-assured support that cannot easily be replicated through automated systems.

AI can support request routing, translation, summarisation and knowledge discovery, but trust, contextual understanding, disciplinary expertise and responsibility remain fundamentally human.

We discuss the organisational principles of the network, lessons learned from its development, and opportunities for extending this approach to other research domains. Our experience suggests that distributed human expertise, connected through shared practices, provides a sustainable model for research data support in increasingly complex



research environments and demonstrates how people remain at the centre of effective data stewardship.

**Centralized or Decentralized? A Comparison of Two Data Stewardship Models at Swiss Universities**

In 2021, Switzerland adopted the National Open Research Data (ORD) Strategy to promote FAIR data practices and strengthen the services and infrastructures supporting research data management. As part of this initiative, swissuniversities has funded the implementation and expansion of data stewardship at Swiss higher education institutions since 2023, while allowing institutions to develop models suited to their local contexts.

This paper compares the implementation of two contrasting data stewardship models at comprehensive Swiss universities of similar size. The University of Basel coordinates a decentralized network of approximately 35 departmental data stewards, supported by a small central management team of ORD experts. The University of Bern employs five full-time, centralized, domain-specific data stewards serving the entire university. Despite offering similar services, including research data management (RDM) consulting, training, and community engagement, the two models differ substantially in their organizational structures, governance, and sustainability.

The decentralized model benefits from strong integration within research communities, local trust, and flexible use of existing expertise, but faces challenges related to coordination, competing responsibilities, and sustainability of services. The centralized model enables strategic planning, specialized expertise, consistent service delivery, and stronger integration into institutional governance, while facing greater challenges in community engagement and visibility within research groups.

Both institutions report common obstacles, including communicating available services, supporting researchers with diverse levels of ORD expertise, and navigating technical and institutional constraints. Shared success factors include strong institutional support, clear policies, and recruiting data stewards with interdisciplinary research experience and excellent communication skills.

By comparing these two approaches, we identify lessons learned, highlight the strengths and limitations of centralized and decentralized data stewardship, and discuss how these insights can support the sustainable development of ORD services across the Swiss higher education landscape.




**Centring the Community Through Minimal Computing for Sustainable Digital Cultural Heritage**

In the cultural heritage sector, digital preservation is increasingly hindered by a lack of sustainable infrastructure, funding constraints, and rapidly outpaced technical skills. While centralised state repositories offer long-term archiving, they often demand operational resources that small-to-medium organisations lack, due to constraints such as lack of funding and understaffing. This compromises community agency and local ownership, making it mandatory to rely on external professional figures and feeding an unvirtuous system.

To address these systemic inequities, this lightning talk introduces an ongoing two-year interdisciplinary project led by a Higher Education institution in Ireland. By leveraging minimal computing – prioritising lightweight, low-bandwidth, and energy-efficient frameworks – the project aims to renew the digital stewardship methodology by keeping human-centred workflows at its core.

Grounded in a participatory, mixed-methods Human-Centred Design methodology – combining international expert interviews, national survey of local cultural organisations, and co-design workshops – the project (aligning with UN SDGs 10, 11, 13) explores and tests minimal workflows across textual, visual, and auditory data. We present the project's practical framework and FAIR-aligned outcomes across four core outputs:

- A collection of essays establishing critical frameworks for minimal digital curation;
- An open-access operational toolkit empowering heritage practitioners with autonomous digital skills;
- A sustainable, low-carbon digital encyclopaedia prototype for visual arts (in collaboration with an artist-led partner);
- A decentralised digital storytelling platform for the ethical, secure management of sensitive personal recovery narratives (in collaboration with a community recovery service partner).

By demonstrating how community-driven workflows can be practically implemented in real-world settings, the project aims to offer a scalable, environmentally conscious, and socially responsible model for independent digital stewardship. The accompanying extended abstract delves deeper into the project's empirical findings, technical architecture, and the FAIR-alignment protocols of each output, providing a comprehensive roadmap for implementation.




**Check-My-Dataset: A Pre-Submission Tool to Improve FAIR Data Quality and Curation Workflows**

Research data repositories increasingly need to balance efficient curation workflows with supporting researchers in producing high-quality, FAIR datasets. At KU Leuven, the research data repository RDR, built on Dataverse and CoreTrustSeal certified, has seen a continuous rise in dataset submissions. This growth has amplified the need for scalable approaches that improve metadata quality while supporting researchers efficiently.

To address this, the RDR team initially developed an open-source review dashboard that streamlines curation by combining workflow management with rule-based automated checks. While these checks improved consistency and reduced reviewer workload, recurring issues revealed that many problems could be addressed earlier in the process. In response, we developed the 'Check-my-dataset' tool, a self-assessment application that enables researchers to evaluate and improve draft datasets before the submission.

The tool builds on the same validation logic used in the review dashboard but reframes automated checks as actionable, user-oriented feedback. It performs a structured analysis of metadata completeness, licensing, documentation (e.g. README file), and dataset characteristics such as size or file structure. Beyond flagging issues, the tool provides contextual guidance and FAIR-oriented recommendations.

This talk will focus on how pre-submission self-assessment reshapes the curation process. The tool identifies common issues earlier, reduces the need for multiple review iterations, and clarifies expectations for researchers. At the same time, it reinforces the complementary roles of automation and human expertise: while automated checks enhance efficiency and consistency, human curation remains essential for contextual interpretation and nuanced decision-making.

We further reflect on ongoing efforts to integrate this functionality more deeply into the Dataverse ecosystem, moving towards a more embedded component of the submission and review workflow.

By sharing design choices, implementation insights, and early observations on workflow impact, we show how the tool supports empower researchers and curators, while improving dataset quality and streamlining the publication process.




**CuratioCuration for Reusability or for Shareability? Insights From Institutional Data Stewardship in the Agrifood Sector for Reusability or for Shareability?**
This abstract reflects on six years of managing the institutional research data repository of IRTA, specializing in the agrifood sector, integrated within the Consortied Dataverse CORA.RDR. The repository hosts highly heterogeneous data, where tabular datasets, images, geospatial files, and code constitute the largest volume of published data.

Initial assessments revealed a significant gap between researcher awareness and the actual FAIRness of datasets prior to curation. The vast majority of submitted datasets required intensive professional curation to comply with FAIR principles. Common issues included a complete lack of README files or contextual documentation, poor dataset structures, inadequate variable names, and unstandardized missing value encodings. Furthermore, files were predominantly deposited in proprietary formats, particularly Microsoft Excel, frequently containing multiple tables within a single workbook that demanded manual extraction, reorganization, and conversion into open formats like CSV.

To address these human and technical bottlenecks, the RDM service implemented a capacity-building strategy structured around a front-office and back-office model piloted during 2024–2025. This comprehensive program combined modular training, targeted data literacy sprints, hands-on datathons, and personalized support to elevate data quality before repository submission. To further optimize workflows, plans are underway to deploy a pre-ingest assessment tool enabling researchers to self-evaluate compliance with curation criteria prior to deposit.

Encouragingly, this human-centric investment has driven a positive trend, improving initial submissions and curation efficiency. However, a persistent challenge remains. After six years of practice, we argue that while datasets become discoverable, accessible, and shareable, true interoperability lags behind due to locally defined variables and project-specific structures. This systemic deficit prompts a critical question: are current curation efforts primarily enabling data sharing rather than maximizing genuine disciplinary reuse?

**Data Professionals as “Safe People” – Fostering Competencies in Handling Sensitive Data**
Curating and providing access to sensitive and restricted research data requires the implementation of safeguards to ensure that the data are handled in an appropriate way. Recent developments in the data curation sphere, such as the use of AI-based tools, pose new challenges for governance and technical infrastructures, make the importance of such safeguards all the more evident. Against this background, it is crucial to ensure continuous learning such that researchers and data professionals (such as data curators) are up-to-

speed to be able to handle sensitive data appropriately. To help individuals tackle these problems confidently, we are developing a modular online training program. This training aims at preparing researchers and data professionals to work with sensitive data in a safe, legal and ethical way. The asynchronous approach of the training is scalable and allows for flexible training of people across different countries, services and academic domains. The training covers modules on topics such as the responsibilities of working with sensitive data, legal basics, data ethics, and statistical disclosure risk.

In this lighting talk, we first introduce the objectives and structure of the training. We then define relevant competencies for data professionals working with sensitive data. This definition is synthesized from an established competency framework, a related learning objective matrix, and expertise from professionals working in related fields. Finally, we outline how we translate this inventory of learning outcomes into a modular course structure.

This talk is of interest to all professionals who work with sensitive data themselves (such as data stewards and curators), who manage data professionals or their working environments or who are involved in research data infrastructures in other ways. Thus, this talk contributes a timely and novel training approach for data professionals and thereby fosters the safe, legal and ethical curation of sensitive data.

**Data Stewards @ WUR – Bridging the Gap Between Research and Support**
The Wageningen University & Research (WUR) Data Stewards community is unique in its kind. The WUR Data Steward, being embedded in a research group, offers domain-specific support and knowledge to researchers and PhD candidates next to their role as a researcher. Over the past 8 years, this community has grown to over 170 Data Stewards and 5 coordinating Data Stewards, playing a crucial role in the transition towards FAIR data and Open Science, supporting efficient, reproducible, and sustainable research data management. What have we learned from building and implementing the Data Stewards community? In this talk, we explore the impacts of an embedded network of data stewards, then discuss challenges, lessons learned, and future directions.

**Developing RDM Support Through User Needs and International Benchmarking: Lessons from the Swedish University of Agricultural Sciences**
Research data and environmental monitoring data are increasingly recognised as strategic assets within higher education. At the same time, growing requirements related to FAIR data, open science, legal compliance, and long-term preservation place increasing demands on coordinated and sustainable research data management (RDM) support.




To inform the future development of Data Management Support (DMS) at the Swedish University of Agricultural Sciences (SLU), a strategic vision process was initiated, combining an internal needs assessment with an international benchmarking study. The aim was to identify current and future support needs, understand emerging trends, and provide an evidence-based foundation for a shared vision for data management support.

The internal study, based on interviews with researchers and environmental analysts, highlighted the importance of well-managed data for research quality, reuse, and long-term accessibility. Participants also described challenges related to administrative requirements, legal uncertainty, metadata creation, data publication, storage costs, and the management of large and complex datasets. They consistently requested more practical, user-oriented support, including guidance, examples, workshops, and access to expertise. A recurring concern was the dependence on individual staff members for critical data-related knowledge.

The benchmarking study examined data management support services at ten Swedish and eight international universities and research organisations. Results showed a rapidly evolving field characterised by small but expanding support units, extensive cross-functional collaboration, and growing demand. Support for data management planning, storage, sharing, and publication was generally well developed, while information classification, advanced data processing support, and hands-on technical assistance remained less mature. Major future drivers included increasing demand, AI developments, FAIR requirements, and resource constraints.

Together, the studies demonstrate the value of combining user-centred needs assessment with international benchmarking when developing RDM services and highlight the strategic role of data management support in research quality, environmental monitoring, innovation, and long-term institutional knowledge management.

**Developing RDM Workflows and Guidelines for Domain-Specific Data Types Together with Researchers**

Research Data Management (RDM) support services often face the challenge of balancing general institutional requirements with the highly specific needs of different research communities and data types. To address this challenge, the Open Science Team at the University of Basel launched the project ‘Advancement of the Data Stewardship Program and Research Data Management Support’1 as part of swissuniversities’ Programme Open Science II2.




The project aims to support researchers in implementing the university's data and information classification within their daily RDM practices. Between July 2025 and June 2026, three subject-specific use cases were developed collaboratively with researchers from Biomedical Ethics, Nursing Science, and Pharmaceutical Sciences. The use cases focus on qualitative interview and focus group data, routine healthcare data, and digital survey data.

The project follows a participatory and iterative approach involving 3-5 researchers per use case. Two workshops were conducted for each use case to analyse existing workflows, gaps, challenges, and disciplinary requirements. Based on these findings, two types of documents were developed: First, a concise workflow that provides step-by-step visualizations of data management processes across the research data lifecycle. Second, detailed RDM guidelines that provide in-depth explanations of the workflow steps, including practical recommendations, tools, and references to further resources.

To encourage reuse and adaptation in other fields, the documents were supplemented with adaptation suggestions and lessons learned, made available to all university members, and published on Zenodo3.

This contribution presents a transferable, participatory methodology for developing domain-specific RDM workflows together with researchers, as well as experiences from the collaborative use case process. It further discusses how adaptable workflow templates and guidelines can strengthen data stewardship services, facilitate the implementation of institutional policies, and foster sustainable and FAIR research data practices.

1 https://doi.org/10.5281/zenodo.16313537
2 https://www.swissuniversities.ch/en/topics/open-science/open-science-programme
3 https://zenodo.org/communities/workflows-and-guidelines-for-research-data/

**Did You Ask for a Free Delivery? Ensuring FAIR Datasets from Data Management Planning to Data Publication**

The implementation of the Portuguese National Program for Open Science and Open Research Data (PNCADAI) in 2025 opened the door to a new range of opportunities regarding the umbrella of research data: new national services – a platform for Data Management Plans (DMPs) , a service for active research data and a research data repository – as well as RDM community improvement and the integration with the national and international science ecosystem.




Within the new services and integrations, one activity stood out for its importance and structural significance: the integration between the national platform for creating and managing DMPs (POLEN Blueprint) and the national research data repository (POLEN DataHub).

With this integration, the expectation is that DMPs created in the POLEN Blueprint can be published in the POLEN DataHub research data repository, enabling three new prospects: 1) one place, within the Portuguese national RDM ecosystem, to deposit and publish DMPs, promoting future customizations needed; 2) the connection between DMPs and the datasets deposited in the repository giving a better perspective and context from the planning to the publish and preservation phases; and 3) the FAIRness evaluation of the deposited data (in each DMPs are incorporated) within the data repository, through the creation of a FAIRness Reference Model, supporting standards, protocols and – ultimately – tools for FAIR evaluation of data, within the OSTrails European project in which the Portuguese POLEN Blueprint and DataHub platforms are part of.

**DMP Bridge: An Open-Source Python Pipeline Using LLMs to Turn Static DMP PDFs Into Machine-Actionable JSON maDMPs**

Data Management Plans (DMPs) are moving towards more machine-actionability, as funders, services, and organizations adopt support for machine-actionable data management plans (maDMPs) that allow automations, integrations, and updates over time. The Research Data Alliance (RDA) has created a common standard format for maDMPs, helping ensure standardization across services and tools, and now the API common standard, making it easier for services in the research data ecosystem to be interoperable with each other.

However, many existing DMPs are only in the format of static PDFs that don't allow taking advantage of these newly available APIs and maDMP advances. Additionally, DMPs are created from many different funder templates, which are constantly evolving over time, which makes it hard to create tools that can work with DMPs from any source within the RDA common standard framework.

To address both of these issues, we developed an open-source, Python-based pipeline that converts DMP PDFs from any funder format into a JSON format that is compatible with the RDA common standard.

This talk will discuss the development and testing of DMP Bridge with a wide range of DMPs from different sources and its accuracy in extraction relative to human annotators. It will discuss the strengths and weaknesses of such a pipeline for turning static PDFs into JSON

format and demonstrate potential use cases (including for ensuring the accessibility of DMP information). We will review learnings and pitfalls from this project that may aid others in similar development initiatives, such as which LLMs performed best at the extraction and how we decided on annotation schema and other aspects for human evaluation to compare to the LLM outputs.

**Documenting and Curating Born-Digital Culture: Early Online Exhibitions and the Preservation of Net Art in Portugal**

This talk presents research on the online curation of born-digital objects in Portugal and examines how it responded to the emerging challenges of digital culture. The study analyses three projects created specifically for the Internet between 2000 and 2005 — Go to Frisco (2000), Ligações_Links_Liaisons (2001), and Online Portuguese Netart 1997|2004 (2005) —developed by artists and curators during a period that preceded the full integration of digital technologies into Portuguese society. These projects comprise two collective exhibitions and one interdisciplinary initiative that sought to foster networked artistic creation, present works conceived for the digital environment, and explore the possibilities of navigation, interaction, and communication offered by the emerging web. Beyond their artistic and historical significance, they represent important examples of curatorial practices that challenged dominant institutional models and experimented with new forms of cultural mediation in digital spaces. The analysis of these case studies demonstrates how artists and curators addressed key questions surrounding net art, particularly its presentation both online and within physical exhibition contexts. Through the testing of hybrid exhibition formats, they sought to move beyond the categorisation of net art as merely “art and computation”. These projects also reveal collaborative approaches and sharing-based models that anticipated dynamics that would later become characteristic of contemporary digital culture. From the perspective of digital curation and digital heritage preservation, mapping these initiatives provides insight into how early online curatorial practices in Portugal developed in parallel with international debates on digital culture and networked artistic production. At the same time, it highlights the importance of documenting, preserving, and recontextualising born-digital projects, which are often vulnerable to technological obsolescence. The talk thus contributes to ongoing discussions on the curation of digital objects as an essential practice for ensuring the accessibility, interpretation, and long-term memory of cultural expressions created on and for the Internet.

**Enabling FAIR by Design Data Infrastructure for Sustainable Biofoundries**

Biofoundries conduct multidisciplinary biotechnology research to deliver meaningful impact on societal challenges. However, transforming large-scale biological data into

actionable insights is often hindered by data silos and inefficient data management environments, which compromise data traceability and reusability.

Data standardization for Biofoundries' innovation processes is critical for accelerating research development and fostering a FAIRer data lifecycle for delivering microbial cell factories and bio-based solutions. To address these challenges, the Informatics Platform at the Novo Nordisk Foundation Biotechnology Research Institute for the Green Transition at the Technical University of Denmark (BRIGHT) established a FAIR-by-design infrastructure governing the data lifecycle from acquisition to analytics. We implemented this through an API-driven ecosystem that integrates six core components: a Lab Information Management System (LIMS), an in-house Data Catalogue, Data Provenance, Azure Data Lake, Nextflow pipelines, and the Seqera interface.

The Informatics Platform is strategically organized into three functional units: data science, software engineering, and a standards office. These teams provide the tools for registering and standardizing experimental data, as well as analysing it while ensuring robust data management and traceability. Furthermore, a network of "satellite enablers" (LIMS super users, data ambassadors in research groups) together with the standard office drive data literacy and maintain agile feedback loops to ensure effective data stewardship.

Going forward, the Informatics Platform aims to continue developing open-source tools and standards that can be used by the Biofoundry community and demonstrating how data can contribute to the green transition.

**Establishing and Consolidating a Data Stewards Network: A Community-Centred Approach to Research Data Stewardship**

The Portuguese Data Stewards Network was created to connect professionals involved in managing, supporting, curating, and enabling the reuse of research data. Launched in 2025, it provides a national space for collaboration, learning, and exchange among those working in research data management and open science. As both a professional network and a community of practice, it brings together data stewards and related professionals from universities, research institutes, competence centres, libraries, and research support services.

In its first years, the network grew quickly, bringing together more than one hundred members from dozens of institutions across the country. Monthly meetings, a mailing list, a website, and online community spaces helped establish regular channels for communication and peer exchange. A governance model, including a Steering Board

representing diverse institutions and areas of expertise, was implemented to support shared responsibilities and long-term sustainability.

One of the network's main early achievements was the launch of a national survey to characterise professionals involved in research data management support. The survey provided the first comprehensive picture of this emerging community in Portugal, identifying roles, responsibilities, training needs, and institutional challenges. Its results highlighted the need for clearer competency profiles, career pathways, and professional recognition, leading to the creation of a working group on data steward profiles and competencies.

The network also contributed to national capacity building through structured training initiatives, including a Data Steward Training Bootcamp and a Train-the-Trainer Workshop. These activities strengthened practical skills, supported institutional knowledge transfer, and marked an important step towards a sustainable training ecosystem for data stewardship.

These early results show the importance of supporting the people and communities behind research data management, FAIR data practices, and digital curation. By fostering collaboration, skills development, and professional recognition, the Portuguese Data Stewards Network strengthens national data stewardship capacity and aligns with IDCC27.

**Hardly a Catwalk: Why Business Models Deserve More of a Spotlight in the Realm of RDM Services**

This submission argues that the Business Model Canvas (BMC), a well-established strategic tool in commercial product and service development, also offers a powerful but underused lens in the context of digital curation: for designing sustainable research data management (RDM) services. It draws attention to core questions needed to make any service, from consultation via curation to compute, user oriented in that it sharpens the role of value contribution. This in turn helps determine how resources for operating a service sustainably could be mobilised.

The talk draws on practical experiences from Germany's National Research Data Infrastructure (NFDI), a nationwide, bottom-up initiative funded by the federal and state governments to systematically manage, preserve, and make accessible research data across disciplines. NFDI organises 26 disciplinary and methodological consortia plus the cross-cutting initiative Base4NFDI that co-designs user-friendly basic services to support all communities.

The talk presents three case vignettes from different contexts that illustrate how applying the BMC helps clarify value propositions, stakeholder constellations, and cost–revenue logics for RDM services under development. NFDI4Culture's Service Center for LIDO (Lightweight Information Describing Objects, a data delivery format) helps institutions network their collection data by offering an integrated service package with training, personalized consultation, and best-practice guidance to improve data quality, and advance the LIDO format nationally and internationally. NFDI4Society's data forum Forum4MICA, fosters consultation, peer support, and circulation of expert knowledge for complex, often highly sensitive, social science data. Base4NFDI's DMP4NFDI shows how a cross-discipline central, multi-tenant platform for data and software management plans coordinates helps researchers and institutions through template development, consulting, and hosting.

In the context of RDM infrastructures, business models can help to show, how services can be sustainably maintained and evolved under constrained public research budgets. The tools that help do this, deserve more of a spotlight.

**How Thick Can You Get? A Discussion on Metadata Awareness of Data Depositors and Reusers**

Creating data archives where users actively engage with data professionals, ask questions, and read documentation is challenging. Therefore, we must understand our users better. During data deposit, we often encounter only the absolute minimum for data documentation (metadata, codebooks, data dictionaries). However, when we explain the necessity, depositors will add more information where directed. This suggests a general lack of awareness around data documentation, not laziness or carelessness. A new State Archives of Belgium project is aimed at measuring the level of understanding of metadata for data professionals, researchers, and the general public. While we expect a sharp decline in metadata awareness from data professionals to the general public, we are curious how specialised our knowledge is. Here we will discuss this survey, areas of investigation, and preliminary results.

To parallel this work, we also present a review of UK Data Service data queries. The UKDS provides detailed metadata records and documentation for all data available through our data catalogue. Guidance on catalogue use is available to all researchers, though only UKDS SecureLab researchers are required to attend training including a detailed review of the documentation. Despite the presence of this guidance, the UKDS User Support Team routinely answers queries from researchers which could have been answered from the available metadata and documentation. We will discuss the common questions, data set and variable patterns, and presentations of answers.




Thus, we will discuss the two sides of metadata awareness, data depositors who must understand why they need to document their data extensively, and data reusers who need to understand how to utilise the metadata. While we assume that such awareness is evident, we argue that the majority of data depositors and reusers are lost in the maze of documentation and we need to change our engagement to address this awareness gap.

**How to Curate a Million Websites: Arquivo.pt's Curatorial Approach**
This presentation describes a curation case study involving websites preserved by Arquivo.pt (the Portuguese Web Archive service), which relied on manual work and human verification by the digital curator to create lists of websites organized by topic.

Through this case study, we seek to address questions such as: a) Are artificial intelligence services ready to generate the datasets we want? b) To what extent is it possible to automate curation tasks?

Thus, we propose to follow these steps: 1) We begin by explaining what the Arquivo.pt service is; 2) We identify the datasets that have been curated—that is, lists of websites; 3) We compile the criteria for creating the lists and the manual and human-based techniques used; 4) We present three examples of prompts we tested to produce a thematic list of websites, identifying their limitations and potential; 5) We conclude by considering the role of human curation and the potential of automation and AI to support the curator's work.

**Human Centered Metadata Creation and Gen-AI as a Tool**
For digital collections of diverse subjects with cultural, and historical significance; using person-first language as a framework for description is necessary. These descriptive practices, require a human at the centre. With the advancement of Gen-AI and institutional push for the adoption of AI tools, new descriptive workflows are being created. For human centered description of digital materials, ethical, practical, and efficient Gen-AI integration is possible. Responsible AI workflows can be achieved with careful guidelines for inclusive description, institutional privacy agreements, and transparency at the point of access for users. This lightning talk presentation will share a use-case about the integration of Microsoft Co-pilot with metadata creation and person-first description practices. The lightning talk will show how inclusive metadata creation was facilitated using Gen-AI and how the workflow was shaped by local best practices and policy.

**Human Curation in AI-Driven Metadata Enrichment**
AI is often presented as a way to produce archival metadata faster, cheaper, and more scalable. But research on named-entity recognition, genre classification, and topic

modelling has long shown that quality depends on human labour before the model begins: annotators draft guidelines, define categories, create gold standards, and resolve disagreements. As large language models become more capable, it is tempting to imagine that this stage can now be skipped.

In Kroika, we tested that assumption on one of our projects – Russian Independent Media Archive. We asked several LLMs, including reasoning models, to enrich 2,000 documents (in our case – media articles) with seven kinds of metadata: genres, topics, people, organizations, locations, events, and summaries. We deliberately avoided a formal annotation stage to see how far current models could go on their own.

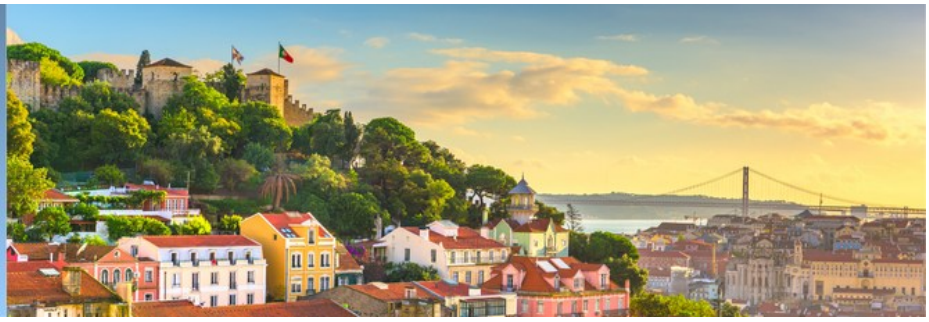
The experiment showed that human curatorial work does not disappear when annotation is removed. It simply moves elsewhere. Before processing could begin, people still had to explain the collection to the model, decide which metadata mattered for the archive and its users, define blurred boundaries between categories, and connect those choices to existing archival standards. Evaluation also required human expertise: assessors needed clear criteria for what counted as a good answer, and they had to apply contextual judgement the model could not provide.

The models were fluent, confident, and often useful. Their errors, though, were patterned: invented entities, missed mentions, venues treated as locations, laws treated as organizations, and genre labels that ignored how journalistic conventions change across outlets and decades. Even the strongest models needed curatorial scaffolding before the task and human assessment after it.

This lightning talk asks where the human-AI boundary actually lies in archival enrichment: what can be delegated, what cannot, and why the human work behind metadata keeps reappearing.

**Human Curators: The Vital Value-Add in the Revision Process**

The increased proliferation of both automation and artificial intelligence in the curation process offers undeniable benefits; however, there are significant risks if these tools are used without human understanding and oversight. Here, we will explore the relationship between these tools and human curation, and evidence why we believe that people have a crucial and inalienable role to play in adding real value to the process. This value comes from a human's ability to contextualize, extrapolate, and empathize better than a machine. Three key areas we will discuss are nuanced situational analysis, the opportunity to educate users, and the opportunity to learn from users and incorporate feedback into the process.



The solutions we develop and implement are informed by a variety of data sources specific to our generalist data publication platform. These include lessons learned from individual cases, analysis of monthly performance reporting, analysis of user feedback from monthly surveys, and ongoing analysis of user behaviour from repeat submitters.

Through engagement with both datasets and their authors, curators are uniquely placed to understand both the immediate challenges and opportunities of data publishing. The automation of standard, repetitive tasks frees up time for these human curators, allowing them to add real value to the curation process.

The field of AI will continue to improve, and an efficient approach to curation demands the exploration and balanced advocacy for and use of AI and automation. At this time, however, there are no accessible AI tools which are able to perform the entire curation process to the required high standards that uphold the key values inherent to curation. Until such a time that technology advances, human curation remains the leading and vital component in ensuring optimal author interaction, and in ensuring the publication of the clearest and most reusable version of a submitted dataset.

**Improving Data Curation Workflow and Support Materials Through User Engagement: The DataRepositóriUM Use Case**

DataRepositóriUM, the institutional research data repository of the University of Minho, is implemented using Dataverse and aligned with Open Science and FAIR principles. Recent repository development has been strongly influenced by the CoreTrustSeal (CTS) certification process and ISO 9001:2015 quality management requirements, leading to the formalisation of curation workflows, clearer quality assurance procedures, and enhanced curator–researcher interactions. This contribution presents the curation workflows implemented in DataRepositóriUM and evaluates them through user feedback collected to support continuous service improvement.

The repository operates through distinct user roles with different levels of autonomy and curation responsibility. To assess users' experience of these workflows, an online survey was distributed to users who had previously deposited datasets in DataRepositóriUM, to evaluate their user experience, perceived impact of curation activities, quality of helpdesk support, and preferences regarding future deposit models.

Results indicate a generally positive user experience. Over half of respondents considered the deposit process easy or very easy, while 60% found the workflow clear or very clear. The repository remains strongly oriented towards mediated deposit, with 56.3% reporting that

their datasets are reviewed before publication. Most respondents valued curation support, with 93.7% considering curator review useful, and helpdesk services received positive ratings for both response time and guidance quality. Survey findings also identified areas requiring improvement, particularly licensing guidance, metadata support, documentation and file organisation, and clearer explanations of the curation workflow.

The results highlight the importance of combining structured curation workflows with personalised support services. User feedback is being used to refine guidance materials and workflow communication, supporting greater user autonomy while maintaining repository quality assurance standards.

**Making Everyday Heritage FAIR: Co-Creating Metadata with Its Communities**
Efforts to make cultural heritage data FAIR (Findable, Accessible, Interoperable, and Reusable) are predominantly expert-driven and centred on assets held by formal memory institutions, often overlooking the heritage that individuals and communities themselves value, which leads to interoperable datasets that fundamentally reflect archival priorities rather than the diversity of lived cultural experience. This paper addresses this limitation by proposing a community-driven framework for creating FAIR metadata for everyday heritage: the places, objects, practices, and memories that are culturally significant to communities but frequently absent from institutional custodianship. We argue for a reframing of memory institutions as co-authors and stewards of collectively produced knowledge rather than sole authorities over metadata. To support this shift, we introduce a new process-oriented framework that defines roles, responsibilities, and a five-stage lifecycle for metadata creation, enabling the coexistence of community folksonomies and institutional taxonomies without erasing diverse perspectives. The framework is operationalised through a metadata model compatible with the European Metadata Model, designed to support plural interpretations, multilingual expression, and incremental enrichment while remaining interoperable with large-scale heritage infrastructures. To validate its feasibility, we conducted a multilingual proof-of-concept study in which participants documented tangible and intangible manifestations of everyday heritage using a structured yet flexible metadata schema. The results demonstrate both the diversity of community-valued heritage and the practical viability of community-led metadata creation. By extending FAIR principles beyond institutional collections, this work contributes a scalable, standards-aligned approach to participatory heritage documentation, supporting more inclusive, equitable, and plural cultural memory ecosystems.




**National Catalogue of Repositories (NKR): Building a National Discovery Infrastructure for FAIR Digital Objects in the Czech Republic**

The National Catalogue of Repositories (NKR) is a central information infrastructure being developed by the National Library of Technology (NTK) in Prague as part of the Czech National Repository Platform (NRP). It addresses the increasing fragmentation of the Czech research landscape, where research outputs are dispersed across diverse institutional and thematic repositories. Unlike a metadata harvester, the NKR functions as a curated registry that maintains authoritative information about the repositories themselves, including their technical platforms, metadata schemas, persistent identifier support, and FAIR maturity.

The primary objective of the NKR is to provide transparency and support for various stakeholders. Data stewards utilize the catalogue to recommend suitable repositories for research data, while researchers can more easily discover outputs. For repository managers and funders, the system offers a way to monitor compliance with national and international standards and improve the visibility of services.

Technically, the NKR is built on two core open-source components: Cordra, which serves as the digital object repository, and Skosmos, which manages controlled vocabularies. The system is integrated with the e-INFRA CZ Authentication and Authorisation Infrastructure (AAI) for secure access.

The NKR fills a critical gap by providing national-level context that international registries, such as re3data, may lack-specifically regarding institutional publication repositories and specialized digital collections. Ultimately, the project aims to make repository metadata itself FAIR, ensuring that the Czech repository ecosystem is machine-readable, interoperable, and fully integrated into the wider European Open Science infrastructure.

**“Open” Does Not Mean “Accessible”: Accessibility as a Missing Standard in Research Data Management**

Data accessibility is a key prerequisite for equal participation in science and is therefore an essential component of Open Science. However, accessibility within the FAIR framework has so far only been addressed from a technical perspective. Drawing on the POUR principles, this presentation provides practical recommendations for the accessible creation, provision, and reuse of research data. Furthermore, the potential and limitations of artificial intelligence for improving the accessibility of research data are discussed. We conclude by presenting an RDM training tool that was designed together with people with disabilities, to ensure that potential barriers are identified and removed at an early stage. With this work, we contribute to the implementation of inclusive science through

infrastructure provision, data literacy and participatory approaches. We argue that openness only becomes truly accessible through barrier-free access.

**People in the Spotlight: Citizen Science Data and Long-Term Preservation**
The theme of this year's IDCC conference acknowledges the crucial role of people in the work of making digital objects FAIR. FAIR digital objects form one of the foundational elements of Open Science. Citizen science is another crucial building block of Open Science, which focuses on involving the public, local communities and stakeholders in the research process, from data collection to defining research questions. Citizen science also centres people and bridges the gap between the scientific community and the general public.

Data generated by citizen science groups have become an increasingly important source for scientists. Also in the humanities we see important citizen science projects focusing on data gathering. Still there are barriers to overcome, both in terms of an inherent lack of trust in citizen-generated data, as well as (in some cases) inconsistent adherence to best practices for data management and preservation. In this presentation these will be discussed in the context of two national citizen science initiatives in the Netherlands in which DANS is involved as the trustworthy digital repository for the long-term preservation of the data.

**RO-Crate as a Bridge Between Communities: Exploring FAIR Digital Objects Across Belgian Federal Scientific Institutions**
Belgium's Federal Scientific Institutions (FSIs) steward a wide range of scientific and cultural heritage resources, including natural history collections, archives, biobanks, heritage science data, conservation records, and other research outputs accumulated through decades of publicly funded activity. While these collections differ substantially in content, scale, and disciplinary traditions, they share a common challenge: ensuring that the knowledge they contain remains understandable and reusable beyond the circumstances in which it was originally created.

Although significant progress has been made in improving the findability and accessibility of research outputs, interoperability and long-term reuse remain more difficult. Data alone is rarely sufficient. Meaningful reuse depends on preserving contextual information, including provenance, methods, documentation, and relationships that enable future users to understand how research outputs were created and how they should be interpreted.




This lightning talk presents an emerging cross-institutional exploration involving several Belgian Federal Scientific Institutions from the natural sciences, archives, and cultural heritage sectors. The initiative investigates whether Research Object Crates (RO-Crate) could provide a practical and lightweight mechanism for improving the portability of research objects while respecting institutional autonomy and disciplinary diversity. Rather than requiring common repositories, infrastructures, or workflows, RO-Crate offers a portable layer through which existing systems can expose and exchange richer descriptions of research outputs.

The exploration examines whether such an approach can operate across institutions with substantially different collections, metadata traditions, and technical environments. Particular attention is given to the role of disciplinary communities in shaping the contextual information that makes digital objects reusable and to the challenge of enabling institutional knowledge to travel alongside digital objects across organisational boundaries.

As an early-stage initiative, the work provides a practical cross-domain perspective on how FAIR Digital Object approaches can support portability, interoperability, and long-term reuse while accommodating the diversity of existing institutional practices.

**Scoping for Sustainability: A Human-Centered Framework for Academic Library Data Services**

Academic libraries have long supported data-intensive research through curation, repository management, and instruction. As demand for these services grows — driven by funder and publisher mandates alongside open science initiatives — libraries must confront the limits of human capacity required to realize the technical goals of the FAIR principles. The growing emphasis on machine-actionability, particularly for AI-ready data and infrastructure, risks marginalizing the people responsible for planning, delivering, and assessing data services. Furthermore, service assessment models that prioritize service growth above all else risk eroding the human competency essential to making offerings sustainable.

In this lightning talk, we present a scoping matrix designed to protect the human element of data services. By mapping service topics against delivery modes, the matrix gives organizations a visual framework for identifying underutilized offerings and aligning goals with actual resource capacity. Rather than expanding services to a breaking point, this tool enables practitioners to prioritize robust, high-impact support. Moreover, this tool is valuable as a process in and of itself: in periods of decreased capacity and resources, articulating what is currently happening, and where the gaps are, is essential for setting

realistic boundaries around services. We argue that consciously scoping our offerings is how we ensure the labour of making data FAIR is recognized, supported, and sustained.

**Social Interoperability in Practice: Reflecting on an Institutional Data Management Partnership After 10 Years**

Within the United States the data management and sharing landscape has been evolving over the past twenty years with growing mandates from federal funding agencies (e.g., the National Institutes of Health, National Science Foundation, etc.), publishers, and disciplinary norms focused on FAIR data and reproducible research. At Duke University two research support units, the Duke University Libraries (DUL) and the Duke Office of Scientific Integrity (DOSI), have been collaborating for the past 10 years to support institutional data management needs. Academic support units bring unique skills and perspectives to engage researchers in this complex environment including knowledge of policies, institutional compliance requirements, data curation best practices, data privacy and ethics, and the list goes on. No one unit can do this work alone, and both formal and informal relationships lay a foundation for effective collaboration. Conceptual models, such as “social interoperability” defined as “the creation and maintenance of working relationships between individuals and organizational units that promote collaboration, communication, and mutual understanding” gives us a general framework to situate this imperative work (Bryant, Dormund, & Lavoie, 2020). In this lightning talk two staff members, one in DUL and one in DOSI, that have been working together consistently for almost a decade, will use a real-world partnership to examine “social interoperability” in practice. We will reflect on how we built and maintained our relationship, provide examples of projects that helped establish a shared understanding of data management, and share lessons learned using an organizational theory lens.

Bryant, R., Dortmund, A., & Lavoie, B. (2020). Social Interoperability in Research Support: Cross-Campus Partnerships and the University Research Enterprise. Dublin, OH: OCLC Research. https://doi.org/10.25333/WYRD-N586

**The Node for Personal and Sensitive Data of the Swiss Research Data Support Network**

Many funding agencies require FAIR and Open Research Data (ORD) to increase the reusability of research results and to make them accessible to the public. This leads to specific concerns for researchers who work with personal data. Even if in principle they are often in favour of the concept of ORD and open science in general, they must adhere to data protection legislations to uphold the interests of the study subjects.



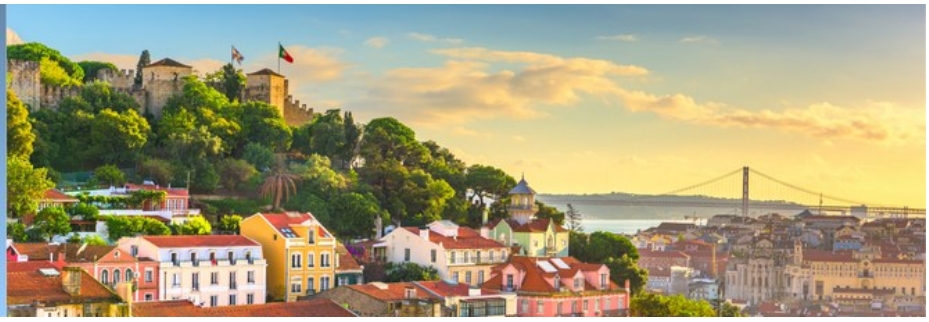

To facilitate transparency while navigating the potential friction between ORD and data protection legislations, and to address topics related to research with personal data the node for personal and sensitive data (PSDN) of the Swiss Research Data Support Network has been established. The PSDN was launched in fall 2025 and currently consists of more than 70 members from 30 Swiss Higher Education Institutions (HEIs) and research centres. To benefit from this broad expertise from different institutions and address their specific needs, node members have the possibility to participate in working groups within the node to increase productivity and develop specific outputs, e.g. software overviews, train-the-trainer inputs or educational resources. Additionally, all node members convene at least biannually in general meetings for an informal exchange, information on the progress of the working groups, or to attend input lectures from internal or external experts. The focus of the node is defined in a solution-oriented way by its working groups and shifts every year. The year of 2026 was dedicated to the identification and documentation of practical tools and workflows for handling personal data in research and collecting information on research data governance practices across SRDSN members of Swiss HEIs.

**The University of Exeter's DRI Journey: Industry Partnerships That Enable Digital Data Stewardship in an Academic Research Environment**

Research data in UK universities has long existed in a fragmented landscape — siloed across local storage devices, unmanaged cloud services, and some centrally managed storage. Often governance frameworks have struggled to keep pace with the volume, velocity, and sensitivity of modern research outputs. At the University of Exeter, we recognised that replacing end-of-life infrastructure was not simply a procurement challenge; it was an opportunity to fundamentally reshape how research data is managed, curated, and made discoverable across the institution.

This talk describes Exeter's Digital Research Infrastructure (DRI) programme: a journey from requirements-gathering and the tendering process through to strategic partnership with industry suppliers. This focused not just on delivering hardware, but a tiered service catalogue, automated lifecycle management, and, critically, the human infrastructure needed to translate technological possibilities into real-world research practice. Central to our vision is the embedding of metadata tagging and data curation workflows directly into the research data lifecycle from the point of project inception. Thus, moving away from retrospective curation as an afterthought and towards more FAIR-aligned stewardship as a default behaviour.

I will describe the critical steps in the process, what we have initialised already, where people remain irreplaceable in this process (humans-in-the-loop), adoption work to embed



new behaviours in staff, and where we intend to take things further. If time allows I may include lessons learned, and what we would do differently next time.

The DRI programme is ultimately not only about storage; it is about social engineering and reshaping and supporting the working habits of the researchers to embrace the FAIR principles from the beginning of every project, increasing the impact of what they do.

**This Is Not a Paradox: Automation for Humanisation in Data Curation**
As the nature of science becomes more collaborative and transdisciplinary, associated datasets grow in both size and complexity and routine ingestion activities consume more of both curator and depositor time. We are developing an improved data curation process for our domain-specific research data repository that manages nationally important datasets of long-term value. In this talk we will present the co-design approach used in this project, share what we have learnt so far and highlight some of the changes we have made to our process and systems.

Our data repository staff help researchers by ensuring that the data they produce are curated appropriately for the long-term, guaranteeing their security, accessibility and reusability for the user community. We curate data and metadata that are trusted, of high quality and machine-readable, which is increasingly important to ensure that they are reusable now, and continue to have impact in the future. As we revamp our curation process, we are ensuring that people remain at its heart and that our high curation standards are retained.

Our objectives are to automate labour-intensive administrative tasks, streamline workflows, automate metadata enhancements where possible, embrace our staff's specialist skills and knowledge, and allow more purposeful communication between staff and depositors.

Our human-centred co-design approach involves iterative feedback sessions with staff and depositors. These sessions are informing the design and development of our revamped curation process and systems through mapping workflows, identifying pain points, reviewing existing tools and discussing opportunities for automation.

The renewed curation process will allow staff to better utilise their data curation skills and knowledge for greater impact. For depositors it means reduced friction and greater agency throughout the process. These changes will contribute to a more transparent and informed FAIR data curation process, encouraging a culture of collaboration between data depositors and curators.




**Toolkit on Legal Issues, Data Protection and Licences**
Open science policies encourage the preservation, sharing and reuse of research data, while data protection law requires researchers to determine under what conditions personal data may be collected, retained, shared and reused. Researchers and data stewards often lack practical guidance that connects these two frameworks and translates them into decisions throughout the research data lifecycle.

This lightning talk presents the Re.Data Toolkit on Legal Issues, Data Protection and Licences, a practical, question-based resource for context-sensitive decision-making in research data management.

The Toolkit was developed through collaboration between Data Protection Officers, ethics committee members, researchers, data stewards and legal experts. This interdisciplinary process addressed recurring tensions between legal requirements, ethical principles, open science objectives and research practice.

A distinctive contribution of the Toolkit is its framework for aligning open science objectives with GDPR requirements. Rather than treating data sharing and data protection as opposing principles, it helps researchers determine when and under what conditions research data may be retained, shared, reused or openly published. A key practical lesson is that these objectives can often be reconciled through layered consent: separate choices regarding participation in the study, future contact and reuse, and the open publication of anonymised data, allowing participants to authorise some operations while declining others.

The Toolkit is organised into sections based on questions about the research project. These questions form a decision pathway that provides users with legal and technical guidance, practical examples, templates, and supplementary resources.

By embedding legal and ethical decisions throughout the research data lifecycle, from data collection and risk assessment to preservation, controlled access, reuse, licensing and erasure, the Toolkit supports legally and ethically informed digital curation. The first version was published in 2025 and is openly available on Zenodo. A revised version and English-language documentation are currently being prepared.

**We PREFER Connected Data: An Ontology for the PREcision FERmentation Community**
Precision fermentation is transforming sustainable biomanufacturing by using engineered microbial organisms to produce food, chemicals, pharmaceuticals, and biofuels. However,

extracting knowledge from large volumes of heterogeneous bioprocess data remains a major challenge due to differences in terminology, metadata, and data structures that make it difficult to reuse these data, and compare across experiments.

To address this challenge, we developed PREFER (PREcision FERmentation ontology), an open-source ontology that provides a shared semantic framework for representing precision fermentation knowledge. Built in alignment with the Basic Formal Ontology (BFO) and connected with established community ontologies, PREFER captures the concepts, processes, relationships, and contextual knowledge that domain experts use throughout the bioprocess lifecycle, making this knowledge explicit, machine-readable, and reusable across labs, software systems, and organisations.

Beyond enabling semantic interoperability, PREFER helps bridge experimental planning, data acquisition, and downstream analysis. PREFER is currently being implemented as the semantic foundation of a biotechnology research institute improving data quality, facilitating reuse, and supporting AI workflows. PREFER is designed as a community ontology, aiming to become a shared semantic standard within the Global Biofoundry Alliance, supporting collaborative digital curation and FAIR bioprocess data across the precision fermentation community.

**When Infrastructure Meets Stewardship: How a Data Storage Project Became a Catalyst for Human-Centred FAIR Data Stewardship**

Over the past years, Hasselt University (UHasselt, Belgium) has strengthened substantially its capacity to support FAIR (Findable, Accessible, Interoperable and Reusable) data practices. These investments include not only dedicated data stewardship support, but also new software solutions and the introduction of the UHasselt Research Drive, an institutional platform for the secure storage and long-term preservation of research data. While such infrastructure is essential, our experience shows that technological solutions alone are not sufficient to achieve sustainable FAIR data implementation.

The implementation of the Research Drive aligned with a broader institutional digital transformation strategy, which also included the migration of research data from Google Workspace to new storage environments. What initially appeared to be a technical infrastructure project quickly evolved into a university-wide change management exercise, requiring researchers to revisit established workflows, data organization practices, retention strategies, and responsibilities related to data governance.

To support this transition, UHasselt's data stewards launched a Research Data Management (RDM) Roadshow, engaging all research groups through tailored

consultations, practical workshops, and a modular RDM Toolkit. These activities aimed not only to facilitate data migration and storage decisions, but also to strengthen awareness and adoption of FAIR and sustainable data management practices.

This lightning talk reflects on the lessons learned from this process and highlights the central role of researchers and data stewards in translating FAIR principles into everyday research practice. Particular attention is given to two use cases from the BIOMED Research Institute: the FIERCELab, working with complex preclinical imaging and stem cell datasets, and the Advanced Optical Microscopy Centre, one of UHasselt's most data-intensive research environments. Together, these cases illustrate how successful FAIR implementation depends on local research cultures, researcher engagement, and tailored support strategies. We argue that infrastructure creates the necessary conditions for FAIR data management, but that people ultimately determine its success.