

Editorial Information

ERCIM News is the magazine of ERCIM. Published quarterly, it reports on joint actions of the ERCIM partners, and aims to reflect the contribution made by ERCIM to the European Community in Information Technology and Applied Mathematics. Through short articles and news items, it provides a forum for the exchange of information between the institutes and also with the wider scientific community. This issue has a circulation of about 2,000 printed copies and is also available online at <https://ercim-news.ercim.eu>.

ERCIM News is published by ERCIM EEIG
BP 93, F-06902 Sophia Antipolis Cedex, France
+33 4 9238 5010, contact@ercim.eu
Director: Dominique Hazaël-Massieux,
ISSN 0926-4981

Contributions

Contributions should be submitted to the local editor of your country

Copyright notice

All authors, as identified in each article, retain copyright of their work. ERCIM News is licensed under a Creative Commons Attribution 4.0 International License (CC-BY).

Advertising

For current advertising rates and conditions, see <https://ercim-news.ercim.eu/> or contact peter.kunz@ercim.eu

ERCIM News online edition

<https://ercim-news.ercim.eu>

Subscription

Subscribe to ERCIM News by sending an email to en-subscriptions@ercim.eu

Editorial Board:

Central editor: Peter Kunz, ERCIM office (peter.kunz@ercim.eu)

Special Theme guest editors:

- Leonardo Candela (CNR-ISTI)
- Roberto Di Cosmo (Inria and University Paris Cité)

Local Editors:

- Ferran Argelaguet, Inria, France (ferran.argelaguet@inria.fr)
- Andras Benczur, SZTAKI, Hungary (benczur@info.ilab.sztaki.hu)
- José Borbinha, Univ. of Technology Lisboa, Portugal (jlb@ist.utl.pt)
- Monica Divitini, NTNU, Norway (divitini@ntnu.no)
- Marie-Claire Fogue, ERCIM/W3C (mcf@w3.org)
- Lida Harami, ICS-FORTH, Greece (lida@ics.forth.gr)
- Athanasios Kalogeras, ISI, Greece (kalogeras@isi.gr)
- Georgia Kapitsaki, Univ. of Cyprus, Cyprus (gkapi@cs.ucy.ac.cy)
- Annette Kik, CWI, The Netherlands (Annette.Kik@cwi.nl)
- Alexander Nouak, Fraunhofer-Gesellschaft, Germany (alexander.nouak@iuk.fraunhofer.de)
- Laura Panizo, University of Malaga (laurapanizo@uma.es)
- Erwin Schoitsch, AIT, Austria (erwin.schoitsch@ait.ac.at)
- Thomas Tamisier, LIST, Luxembourg (thomas.tamisier@list.lu)
- Maurice ter Beek, CNR-ISTI, Italy (maurice.terbeek@isti.cnr.it)

Contents

JOINT ERCIM ACTIONS

4 Message from the President

5 ERCIM “Alain Bensoussan” Fellowship Programme

SPECIAL THEME

Introduction to the Special Theme:

6 Advancing Open Science: Federated Infrastructures and Trustworthy Ecosystems

by the guest editors Leonardo Candela (CNR-ISTI) and Roberto Di Cosmo (Inria and University Paris Cité)

Research assessment and scholarly representation

8 The OpenAIRE Graph: Enabling Open Research Intelligence Using Open Data

by Paolo Manghi (CNR-ISTI)

9 Rethinking Researcher Profiles in the Research Assessment Transition Era: The OpenAIRE Approach

by Stefania Amodeo and Zenia Xenou (OpenAIRE AMKE)

10 BibTexViz: Visualizing Research Productivity with Open Science Data

by Jose-Miguel Horcas (ITIS Software, Universidad de Málaga)

12 A 360° View of Open Science in Europe – Indicators, Narratives, and AI-enhanced Policy Intelligence

by Tereza Szybisty (OpenAIRE Amke)

From FAIR data to AI-ready workflows

14 From Data to Knowledge: Open Pipelines for Collaborative Production Research

by Leon Gorissen and Philipp Brauner (RWTH Aachen University) Germany)

16 Scalable Platforms for Reusing Research Infrastructure Data in Open Science

by Magdalena Brus (EGI Foundation), Ville Tenhunen (EGI Foundation), Gergely Sipos (EGI Foundation), on behalf of the RI-SCALE consortium

ture and situating Open Science developments within the broader research and innovation landscape.

Who benefits and how the Observatory can be used

The Observatory is designed for a broad set of stakeholders, including policymakers, national Open Science coordinators, research organisations, funders, and experts. Typical use cases include:

- Tracking progress over time within a country or across Europe using the policies–practices–impacts structure.
- Comparing thematic areas (e.g., Open Access, data stewardship, research software, citizen science) to identify where enabling conditions exist but adoption lags.
- Supporting policy alignment by browsing and interpreting policy documents.
- Informing policy reform discussions, including responsible research assessment, by connecting monitoring evidence with incentive structures and implementation realities.

As Open Science monitoring evolves globally, initiatives such as the OSMI underline the need for more coordinated, transparent, and co-creation approaches. The EOSC Open Science Observatory contributes to this shift by demonstrating how monitoring can be done with and for community, combining openness, methodological transparency, validation, and FAIR data principles. By integrating indicators, narratives, policy intelligence, and trusted external evidence in an accessible and reusable way, the Observatory offers a model for responsible monitoring beyond Open Science itself. In doing so, it strengthens the evidence base for a more open, inclusive, and sustainable European Research Area [2].

This project has received funding from the European Union’s Horizon Europe framework programme under grant agreement No. 101148217. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Executive Agency. Neither the European Union nor the European Research Executive Agency can be held responsible for them.

Links:

[L1] <https://open-science-monitoring.org/principles/>

[L2] <https://www.eoscobservatory.eu/>

[L3] <https://graph.openaire.eu/>

References:

- [1] G. O’Neill, “Monitoring Framework for National Contributions to EOSC and Open Science”, 2025. DOI: 10.5281/zenodo.14919248.
- [2] T. Szybisty, “Monitoring Open Science across ERA: Early Lessons from the EOSC Open Science Observatory”, 2025. DOI: 10.5281/zenodo.16991651.

Please contact:

Tereza Szybisty
OpenAIRE Amke, Czech Republic
tereza.szybisty@openaire.eu

From Data to Knowledge: Open Pipelines for Collaborative Production Research

by Leon Gorissen, (RWTH Aachen University) and Philipp Brauner (RWTH Aachen University) Germany)

Open Science promises better, faster, and more trustworthy research and benefits arise if data can be shared, aggregated, and reused. We present data-to-knowledge pipelines in an engineering context as a practical way to turn open research data into additional insights and tangible benefits across organizations. Drawing on shared data from industrial robots across organisations, the work illustrates how open infrastructures can enable collaboration at scale.

Open Science has become a strategic priority for the European Commission and other funding agencies, with the aim of making research more transparent, reusable, and collaborative. A key element of this vision is the effective use of open data: not only by publishing datasets, but by enabling data to be combined, interpreted, and transformed into knowledge across disciplines and institutions. We present a framework that addresses this challenge through data-to-knowledge (D2K) pipelines (Figure 1), providing a practical approach to implementing Open Science in a complex, collaborative research cluster focused on the digital transformation of production [1].

The article shows how openly shared research data from different organisations can be systematically transformed into insights that benefit multiple stakeholders. Instead of treating data publication as an end point, the proposed pipelines emphasize continuous processes in which data is produced, shared, enriched with meaning, analysed, and then fed back to inform new experiments, models, or decisions. This approach aligns with the principles of Open Science, particularly reuse, transparency, and reproducibility.

A central concept in the work is the use of Digital Shadows [1]. These extend the idea of Digital Twins by providing context- and task-dependent, purpose-driven representations of real-world entities such as machines, processes, or experiments. Digital Shadows capture both data and its contextual meaning. By standardizing how data is described and linked to real-world objects, they make shared data easier to understand and reuse, even beyond its original application. This is particularly important for interdisciplinary research, where data often needs to move across domain boundaries.

We demonstrate the data-to-knowledge pipelines through an industrial case study involving robotic production systems operated at different sites and by different organisations. Trajectory data from identical robot models, used independently by each organisation, is shared through a common open research data infrastructure (the open source framework Coscine [L1]). Because the data is described in a consistent, reusable way and follows the FAIR principles, it can be easily

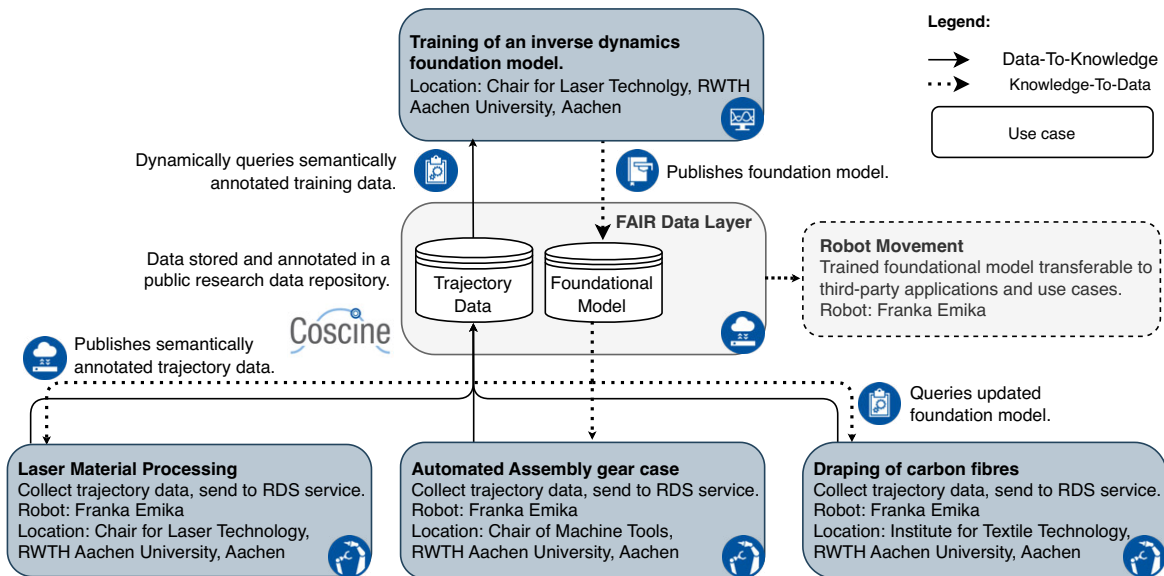


Figure 1: Data-to-knowledge pipeline illustrating how shared production data are transformed into reusable models and insights across organisations.

identified and queried by other processes. This enables an independent second process to query the shared data and train a more robust base model for robot control, benefiting from increasingly diverse trajectory data aggregated across sites without additional effort. The resulting model is equally openly available and improves robotic operation across sites, illustrating how open data can create collective benefits without requiring full centralisation of resources or expertise.

Importantly, the framework respects organizational autonomy while enabling collaboration. Each participant retains control over its own data and systems, while contributing selected data to a shared environment under mutually agreed conditions. This federated approach represents a highly relevant model for Open Science infrastructures, where legal, ethical, or commercial constraints often preclude full data openness. In this setting, data sharing yields reciprocal benefits: contributors and the broader community gain access to a more robust shared base model for robot control, which can in turn enhance local operations while supporting collective advancement. More broadly, the framework enables a “world-wide lab” (WWL) paradigm in which production environments function as distributed experiments, contributing empirical evidence toward a deeper understanding of factors influencing performance, quality, and safety [1,2]. Crucially, participation requires only limited additional effort while offering tangible gains through collaboration.

From an Open Science perspective, the proposed data-to-knowledge pipelines support several key goals. They improve reproducibility by making data findable, accessible, and interoperable, and reusable as data is publicly available and transformations explicit and traceable. They enhance reuse by providing semantic descriptions that allow data to be understood beyond its original project. They also encourage collaboration, as shared pipelines and models create incentives for participants to contribute data and expertise.

The article also looks ahead to future research infrastructures. As scientific data volumes continue to grow, simply storing

open datasets will not be enough. Especially in the domain of resource intensive training of machine learning models, researchers will need shared tools, pipelines, and services that help turn data into knowledge efficiently and responsibly. The presented framework offers a blueprint for such infrastructures, building on open standards and tools, using modular components, and demonstrating real-world applicability.

In summary, the work shows how Open Science can move beyond ideals toward implementation and applicability. By focusing on data-to-knowledge pipelines, it demonstrates how open data can become a living resource, continuously reused, enriched, and transformed, supporting transparent and collaborative research across disciplines and sectors.

Acknowledgements: Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy — EXC-2023 Internet of Production — 390621612.

Link:

[L1] <https://coscine.rwth-aachen.de/>

References:

- [1] P. Brauner et al., “A Computer Science Perspective on Digital Transformation in Production,” *ACM Trans. Internet Things*, vol. 3, no. 2, 2022, doi: 10.1145/3502265.
- [2] M. Behery et al., “Vision Paper: Leveraging Industrial Big Data – Past, Present, and Future of the World Wide Lab,” in *Proc. IEEE Int. Conf. on Big Data*, 2023, doi: 10.1109/BigData59044.2023.10386257.
- [3] L. Gorißen et al., “Demonstrating Data-to-Knowledge Pipelines for Connecting Production Sites in the World Wide Lab,” , *arXiv preprint arXiv:2412.12231*, 2024. [Online]. Available: <https://arxiv.org/pdf/2412.12231>.

Please contact:

Leon Gorißen, Chair for Laser Technology
RWTH Aachen University, Germany
leon.gorissen@ilt.rwth-aachen.de