

OSIRIS as an Open, Flexible CRIS for Distributed Scientific Infrastructures: Insights from its Deployment in the German Network for Bioinformatics Infrastructure (de.NBI)

Martinique Frentrop¹, Sina Barysch², Dominic Koblitz¹, Julius Witte¹, Julia Koblitz¹

1 OSIRIS Solutions GmbH, Germany; 2 EMBL Heidelberg, Germany

Summary

OSIRIS is an open-source Current Research Information System (CRIS) created to support research organisations whose activities include more than traditional publication outputs, e.g conference posters, events or teaching. Its core mission is to offer a configurable, sustainable, and highly adaptable alternative to conventional CRIS solutions that are frequently designed around the needs of universities or institutes.

Scientific consortia and distributed infrastructures often produce software releases, datasets, workflows, training events, and community-building activities that do not fit neatly into rigid, publication-centric systems. OSIRIS was developed precisely to address challenges like this and was successfully implemented at several scientific institutes in Germany.

The German Network for Bioinformatics Infrastructure (de.NBI) provides an ideal example of a modern, distributed research environment where OSIRIS can unfold its full potential. As a nation-wide consortium that coordinates bioinformatics services across eight service centers including 43 institutes, de.NBI requires a CRIS that can adequately capture the diversity of its scientific, technical, and community-oriented contributions. The deployment of OSIRIS within de.NBI demonstrates how a flexible and open-source system can empower a large infrastructure to communicate its outputs effectively, reduce reporting burdens, and improve internal communication and cooperation. This contribution presents OSIRIS through its system perspective and highlights the lessons learned from its adoption in the de.NBI network.

Introduction

The systematic collection, management, and use of research information are becoming increasingly important for research institutions, funding organizations, and political decision-makers. Research information is understood to comprise all structured data that represent scientific activities and their contextual conditions — ranging from publications, projects, and third-party funding to research infrastructures as well as personnel and collaborative structures. It therefore

constitutes metadata from across the entire research lifecycle (cf. Ebert et al. 2016: 7; Fante 2022: 20).

This information forms the basis for internal governance, external evaluation, and science policy decision-making processes. In the context of the Barcelona Declaration (Barcelona Declaration 2024), the value of open research information for Open Science has also been increasingly emphasized (cf. Biesenbender et al. 2019).

To ensure comparability, transparency, and traceability, the Core Dataset for Research (Kerndatensatz Forschung, KDSF) was developed in Germany to serve as a national metadata standard for research information. With the release of version 2.0 in early 2025, the KDSF has undergone substantial further development. The standard, which previously focused primarily on defining aggregated reporting data, has been expanded into a modular core data model. This model enables institutions to record their research activities at the object level—such as projects, persons, or publications—in a structured manner and to flexibly interlink these objects (KFID 2025).

In this context, research information systems (RIS/CRIS) can support the structured collection, management, analysis, and provision of data on research activities for internal governance processes as well as for external reporting (Biesenbender et al. 2019: 148; Azeroual 2021: 145; Angulo-Osorio et al. 2025: 51; Zendulková et al. 2025: 230; Nabavi et al. 2016: 330, 336).

OSIRIS Design Philosophy

OSIRIS is built on the principle of flexible configuration. Rather than imposing a predefined structure, the system relies on a no-code customization toolbox that allows organisations to define their own activity types and categories, projects, workflows and more. This concept ensures that OSIRIS can adapt to scientific practices and organisational needs without requiring code modifications, even when these evolve over time. A second foundational principle is its commitment to openness and interoperability. OSIRIS is entirely open-source and designed to integrate seamlessly with other platforms and institutional systems.

Finally, OSIRIS adopts a deliberately lightweight deployment model. Many research infrastructures lack the extensive administrative or IT resources typically required to operate enterprise-grade CRIS solutions. To address this, OSIRIS relies on a modern yet streamlined technology stack, including containerised deployment options and a flexible MongoDB backend. In addition to the core platform, OSIRIS is complemented by a range of supporting services, including managed hosting, initial system setup with legacy data import, user and administrator training, and finally the Portfolio-addon for a full-fledged public research portal. Together, this combination

lowers barriers to adoption and enables even large, distributed research networks to operate OSIRIS sustainably over the long term.

Research Information in the de.NBI Network

The German Network for Bioinformatics Infrastructure (de.NBI) constitutes a distributed, nationally coordinated research infrastructure that integrates multiple service centers and partner institutions to provide bioinformatics tools, databases, cloud-based computing resources, and training for the life sciences (Tauch et al. 2019). Its activities encompass a broad and growing spectrum of projects, services, and collaborative formats, involving heterogeneous actors, funding frameworks, and scientific outputs. As the network continues to expand in scale and complexity, the amount and diversity of information generated across de.NBI are steadily increasing. At the same time, requirements for transparency, interoperability, and standardized reporting are gaining importance at national and international levels, particularly in the context of data-driven research infrastructures, Open Science, and emerging metadata standards. Against this backdrop, the coordinated and structured handling of research-related information is becoming a central cross-cutting issue for the sustainable development and governance of the de.NBI network.

OSIRIS as a CRIS Framework for the de.NBI Network

The adoption of OSIRIS in de.NBI vividly illustrates how these architectural ideas translate into practice. de.NBI sought a CRIS capable of capturing the full breadth of its contributions, spanning software releases, training events, materials and attendance data, publications, outreach activities, and collaborative projects. OSIRIS allows all these activity types to be modeled without changes to the system's core codebase, demonstrating the strength of its configuration-driven approach.

The configurability also extended to metadata definitions and workflows. Each de.NBI service center had the opportunity to contribute its own requirements, which were then harmonised by the coordinating working group into a shared OSIRIS schema. We are currently in the process of integrating the system into de.NBI's broader digital ecosystem. OSIRIS supports identifier-based metadata enrichment by automatically retrieving publication and researcher information via persistent identifiers (e.g. DOI, ORCID) and is being extended to integrate with training and software infrastructures such as ELIXIR TESS. These features significantly reduce manual effort and enhance the completeness and accuracy of the research information being collected.

A particularly valued capability was OSIRIS's ability to generate public, automatically updated portfolio pages. These pages present the network's publications, software tools, training activities, and other contributions in a transparent and accessible manner. As a result, OSIRIS not only supports internal reporting and evaluation

processes but also enhances the visibility of de.NBI's achievements to national and international stakeholders.

Impact and Lessons Learned

The deployment of OSIRIS within de.NBI highlights how powerful open-source, configurable CRIS platforms can be for distributed infrastructures. The system accommodates the network's heterogeneous outputs and provides a harmonised platform for collecting, validating, and disseminating research information. It enables a significant reduction in reporting overhead while improving metadata quality and consistency. Moreover, it offered a sustainable long-term alternative to commercial systems and provided an environment that researchers and administrators could adopt with minimal friction.

At the same time, the deployment highlighted the critical importance of well-defined governance structures to maintain metadata quality, the need for targeted training across distributed organizations, and the benefits of establishing shared practices early in the implementation process. Governance structures must clearly assign responsibilities for data entry, validation, and curation to ensure consistent metadata standards across all participating units. Training programs are essential not only to familiarize staff with the platform but also to build a shared understanding of research information workflows and reporting requirements. Establishing common practices and protocols early helps to avoid fragmentation, ensures smoother adoption across multiple sites, and fosters a culture of collaboration and data stewardship that extends beyond the technical implementation. These considerations reflect structural challenges inherent in any multi-institutional CRIS rollout rather than limitations of OSIRIS itself.

Outlook and Conclusion

The experiences gained from the de.NBI deployment are already shaping OSIRIS's developmental trajectory. At the same time, the deployment has highlighted that implementing a research information system in a large, distributed infrastructure involves significant organisational hurdles, particularly with respect to coordination across multiple institutes with heterogeneous workflows, governance models, and reporting requirements. Addressing these challenges requires not only technical flexibility but also sustained organisational alignment and clear coordination mechanisms.

Future work therefore includes deeper interoperability with European infrastructures, enhanced support for standards such as the Common European Research Information Format (CERIF; Jörg 2010), expanded analytics and dashboard capabilities, and strengthened mechanisms to empower the community to actively shape and contribute to the system. OSIRIS actively incorporates lessons learned from de.NBI and other installations by engaging with participating institutes,

collecting best practices, and systematically preparing guidance for future deployments. Through these developments, OSIRIS aims to become even more adaptable to the needs of diverse open-science ecosystems, including federated projects and cross-infrastructure research environments.

The deployment of OSIRIS in de.NBI demonstrates how an open, flexible, and sustainable CRIS can empower large, distributed scientific infrastructures to capture and communicate their diverse research outputs. Instead of forcing scientific organisations into predefined structures, OSIRIS adapts to their needs - making it an ideal solution for modern research environments that value transparency, openness, and the ability to grow over time.

Literature

Angulo-Osorio, J. F., Valdivia-Portugal, C. D., & Rosas-Paredes, K. (2025). Enhancing research visibility: A comparative study on the implementation of CRIS systems at Universidad Católica de Santa María and its contrast with other universities. *Publications*, 13(4), Artikel 51.

<https://doi.org/10.3390/publications13040051>

Azeroual, O. (2021). Untersuchungen zur Datenqualität und Nutzerakzeptanz von Forschungsinformationssystemen [Dissertation, Otto-von-Guericke-Universität Magdeburg]. Open Access und Forschungsdaten-Repositorium der Hochschulbibliotheken in Sachsen-Anhalt. <https://doi.org/10.25673/45118>

Barcelona Declaration on Open Research Information. (2024). Kramer, B., Neylon, C., & Waltman, L. <https://doi.org/10.5281/zenodo.10958522>

Biesenbender, S., Petersohn, S., & Thiedig, C. (2019). Using current research information systems (CRIS) to showcase national and institutional research (potential): research information systems in the context of Open Science. *Procedia Computer Science*, 146, 142–155. <https://doi.org/10.1016/j.procs.2019.01.089>

Ebert, B., Tobias, R., Beucke, D., Bliemeister, A., Friedrichsen, E., Heller, L., Herwig, S., Jahn, N., Kreysing, M., & Müller, D. (2016). Forschungsinformationssysteme in Hochschulen und Forschungseinrichtungen. Positionspapier. Version 1.1. DINI AG Forschungsinformationssysteme. <https://doi.org/10.5281/zenodo.45564>

Fante, A. (2022). Publikationsprozesse und Berichtspflichten als Impulsgeber für ein Forschungsinformationssystem. Eine Untersuchung am Beispiel einer nicht-universitären Forschungseinrichtung [unveröffentlichte Masterarbeit]. Technische Hochschule Köln.

Jörg, B. (2010). CERIF: The Common European Research Information Format Model. Data Science Journal, 9, CRIS24-CRIS31. <https://doi.org/10.2481/dsj.CRIS4>

Kommission für Forschungsinformationen in Deutschland. (2024). KDSF 2.0: Flexible und qualitätsgesicherte Forschungsberichterstattung nach dem Baukastenprinzip. Grundsatzpapier der KFiD. <https://doi.org/10.58010/KFID:GP1:2024>

Nabavi, M., Jeffery, K., & Jamali, H. R. (2016). Added value in the context of research information systems. Program, 50(3), 325–339. <https://doi.org/10.1108/PROG-10-2015-0067>

Tauch, A., Al-Dilaimi, A. (2019). Bioinformatics in Germany: toward a national-level infrastructure. Briefings in Bioinformatics, 20(2), 370–374. <https://doi.org/10.1093/bib/bbx040>

Zendulková, D., Azeroual, O., & Krivjanská, A. (2025). From data silos to FAIR research: Interoperability and integration in European CRIS. Digitalization and Management Innovation IV 412, 228–240. <https://doi.org/10.3233/FAIA250723>