

OpenAIRE Graph as an Open Research Information Backbone for CRIS Interoperability and Responsible Evaluation

Angeliki Tzouganatou, Zenia Xenou, Stefania Amodeo, Leonidas Pispiringas, Giulia Malaguarnera, Pedro Principe, Natalia Manola, Paolo Manghi, Andre Vieira

Current Research Information Systems (CRIS) occupy a central position in institutional research information management and are increasingly implicated in national reporting, funder workflows, and research evaluation. Despite long-standing investments in standardisation, (e.g. through CERIF, DataCite, Dublin Core), research information remains fragmented across heterogeneous systems, inconsistently structured, and frequently mediated through proprietary analytics layers. This fragmentation becomes particularly consequential as organisations seek to implement responsible evaluation approaches that require research information to be transparent, auditable, and reusable across different stakeholders and contexts. These requirements align directly with emerging international frameworks for open research information and responsible assessment.

The Barcelona Declaration on Open Research Information emphasises that research information used for decision-making, including evaluation, should be open, transparent, and auditable by those being assessed, thereby reducing systemic dependence on closed infrastructures and opaque methodologies. In parallel, the Coalition for Advancing Research Assessment (CoARA) has consolidated a broad reform agenda in which evaluation systems are expected to be context-sensitive and supported by trustworthy evidence infrastructures rather than black-box metrics. Within this agenda, the CoARA Working Group “Towards Open Infrastructures for Responsible Research Assessment” (OI4RRA) has articulated principles, characteristics, and practical instruments intended to help the community specify what it means, technically and institutionally, for infrastructures to be fit for Responsible Research Assessment (RRA), with emphasis on governance, transparency, operational robustness, as well as ethical and inclusive foundations.

In response to these challenges, this paper presents an open integration framework embedded within the OpenAIRE aggregation ecosystem that integrates institutional CRIS systems to the OpenAIRE Graph. The CRIS integration is incorporated into the broader OpenAIRE data ingestion and harmonisation workflow, alongside repositories, journals, aggregators and other scholarly data sources. The contribution is twofold: (i) it describes the technical architecture and the workflow mechanisms that enable institutional CRIS metadata to be harvested, validated, enriched, and integrated into a scholarly knowledge graph, making research information accessible through open APIs for monitoring, evaluation, and responsible assessment (ii) it concludes with an applied demonstration via MyResearchFolio, a researcher-centric profile service built on the OpenAIRE Graph to support Open Science-aware RRA practices, beyond bibliometric indicators.

1. Background and related work

CRIS have become core institutional systems for managing research information on outputs, projects, funding, organisations, and persons, supporting both internal analytics and external reporting. Interoperability nevertheless remains uneven, reflecting differences in system architectures, local data models, and semantic practices. CERIF (Common European Research Information Format) provides a shared conceptual model intended to support interchange of research information across CRIS platforms. In practice, however, interoperability depends on more than adopting the core model: it requires agreed exchange profiles, controlled vocabularies, and implementation guidance that align how organisations encode entities, roles, and classifications. These elements do not “restrict” CERIF;

they make it implementable in a consistent way across heterogeneous CRIS deployments and enable reliable aggregation and reuse of research information.

1.1 Open Research Information as a prerequisite for a trustworthy evaluation

The Barcelona Declaration places open research information at the centre of transparent research systems and explicitly emphasises that, in research assessment contexts, openness guarantees that those being assessed can access the evidence considered in evaluation, thereby strengthening transparency and accountability. This aligns with RRA agendas: reforming indicators and procedures is difficult to operationalise without infrastructures that support provenance signalling, reproducibility of evidence retrieval, and inspectability of the evidence from which evaluative claims are derived.

1.3 CoARA and the OI4RRA WG framing: from principles to infrastructure requirements

CoARA provides a shared reform umbrella, while the OI4RRA Working Group translates the reform imperative into infrastructure requirements and assessment tools. The OI4RRA framework and principles¹, including the Open Infrastructure Checklist² articulate four assessment dimensions (technical robustness, operational capacity, community-centred governance, and ethical & inclusive foundations), offering a structured lens for judging whether research information infrastructures, including graph layers and CRIS exchange mechanisms, are credible substrates for responsible assessment.

2. OpenAIRE Graph as an open research information backbone

The OpenAIRE Graph serves as a backbone layer that aggregates and exchanges research information across institutional and national CRIS landscapes and enables downstream reuse for discovery, monitoring, analytics, and assessment-relevant workflows. In this architecture, CRIS act as authoritative institutional sources for structured metadata, while the graph layer enables cross-source linking, enrichment, and standardised access points through APIs.

OpenAIRE has developed interoperability guidelines based on the CERIF standard to enable research information exchange between CRIS and other systems³. The guidelines cover the interchange of information about research outputs (publications, patents, datasets, software), projects and funding, research actors (persons, organizations, organization units), events, and instruments or equipment. The OpenAIRE profile of CERIF defined in the OpenAIRE Guidelines for CRIS Managers supports harvesting and importing metadata from CRIS systems.

A key technical consideration concerns the exchange of CERIF data through record-based harvesting infrastructures. While CERIF models research information as a system of interrelated entities, protocols such as OAI-PMH disseminate metadata as discrete records without enforcing relational constraints. Referential integrity therefore constitutes a prerequisite of the CERIF-XML export produced by the CRIS system: all referenced entities and linking objects must be explicitly represented and internally

¹ Manola, N., Tzouganatou, A., Kuchma, I., & CoARA WG on OI4RRA. (2025). Open Infrastructures for Responsible Research Assessment: Principles, Framework, and Checklist for Research Organisations. Zenodo. <https://doi.org/10.5281/zenodo.17881692>

² Tzouganatou, A., & CoARA WG on OI4RRA. (2025). Open Infrastructure Checklist for Research Performing and Research Funding Organisations. Zenodo. <https://doi.org/10.5281/zenodo.17699724>

³ Pablo de Castro, Jochen Schirrwagen, Dimitris Karaiskos, Jan Dvořák, Andrea Bollini, Vasilis Bonis, Nikon Gasparis, Victoria Tsoukala, Paolo Manghi, Pedro Príncipe, Progress in the Implementation of the OpenAIRE Guidelines for CRIS Managers, *Procedia Computer Science*,

consistent within the provided dataset. The OpenAIRE Guidelines define structural and exposure requirements for interoperable CERIF-XML dissemination, including expectations on completeness and identifier consistency. Subsequent validation, harmonisation, and integration into the OpenAIRE Graph are performed within the aggregation workflow. The OpenAIRE guidance addresses this by prescribing how CERIF-XML can be exposed via OAI-PMH as sets of records while maintaining referential integrity so that relationship semantics across entities can be reconstructed downstream. Via the OpenAIRE PROVIDE dashboard, the content gateway service of OpenAIRE, CRIS managers are invited to connect scholarly content with OpenAIRE, allowing to enter both the OpenAIRE and EOSC ecosystems and be accessible by millions of researchers, research institutes and networks, research funders, policy makers and citizens. This provide dashboard offers added value services where CRIS managers interact with OpenAIRE by providing detailed information about the metadata aggregation.

2.1 Data model and access mechanisms

The OpenAIRE Graph provides a unified representation of research entities and their relationships and is *exposed* and *made accessible* through documented public APIs and data snapshots that support programmatic access and query-based retrieval. The Graph documentation describes the main entities, which include "research products", "data sources", "organisations", "projects", and the recently introduced "persons". Entity-specific API endpoints support retrieval and query-based access, enabling transparent and reproducible data access for integration into institutional processes.

2.2 Enrichment and contextualisation under a responsible assessment lens

Scholarly Knowledge Graphs become evaluation-relevant insofar as they support contextualised evidence and mitigate distortions such as misattribution, duplication and incomplete coverage of research activity. The OpenAIRE Graph's relational structure enables this contextualisation, by linking research outputs with projects, organisations, and funders. Where enrichment relies on algorithmic or inferential steps, RRA requires transparency about what is asserted versus what is inferred and how uncertainty is represented and communicated. These expectations reflect the Barcelona Declaration's auditability rationale and align with OI4RRA WG's emphasis on evaluation contexts, community-governed infrastructures. The OpenAIRE Graph addresses this requirement by recording provenance information for the metadata, documenting whether the information was harvested, inferred by OpenAIRE, or linked by a user, and the trust associated with the information.

3. Alignment with Barcelona Declaration and OI4RRA: translating principles into system capabilities

3.1 Openness, accessibility, and auditability

The Barcelona Declaration argues that open research information enables evidence-based decision-making and renders evaluation evidence accessible and auditable by those being assessed. In graph-led infrastructures, auditability can be advanced through provenance signals (source attribution), documentation of transformation and integration steps, and stable, queryable access routes that allow third parties to inspect and reproduce evidence extraction. The availability of documented public APIs is therefore not merely a technical feature; it is a governance-relevant mechanism that underwrites inspectability and supports reproducible evidence retrieval in practice.

3.2 Governance, sustainability, and community-centred stewardship

The OI4RRA framing emphasises that infrastructures must not only function technically but also be governed and sustained in ways that support equity and trust. The Open Infrastructure Checklist provides a structured instrument for assessing whether infrastructures demonstrate robustness, operational capacity, community governance, and ethical foundations, criteria directly relevant when

infrastructures are used as evidence substrates for evaluation. Consistent with this orientation, the Barcelona Declaration explicitly expects supported infrastructures to implement good practices for community governance and sustainability, citing the Principles of Open Scholarly Infrastructure (POSI) as an illustrative benchmark. OpenAIRE has publicly documented its POSI self-assessments.

3.3 Responsible evaluation: separating evidence infrastructure from evaluative judgement

A foundational principle for RRA is to distinguish between infrastructures that provide evidence and decision processes that interpret evidence in context. This separation is consistent with CoARA's reform direction and with OI4RRA WG's emphasis on infrastructures that enable plural, context-sensitive assessment practices rather than imposing a single evaluative logic. Graph infrastructures can support responsible evaluation by enabling multi-dimensional evidence, transparent communication of data completeness and limitations, and the preconditions for correction and contestation, especially where provenance and access mechanisms allow stakeholders to interrogate the basis of evaluative claims.

4. Identifiers and semantics: reducing misattribution risks

In evaluation settings, ambiguity and misattribution (of persons, organisations, and outputs) can materially affect outcomes. Interoperability approaches should therefore prioritise persistent identifiers and semantic alignment through controlled vocabularies as baseline safeguards for trustworthy evidence flows. The Barcelona Declaration explicitly highlights the use of persistent identifiers (e.g., DOI, ORCID, ROR) as a prerequisite for linking information from different sources. The OpenAIRE Graph documentation describes persons and their identifier mapping (including ORCID-derived identifiers) as part of its entity model.

5. Evaluation-relevant capabilities: from metrics to evidence ecosystems

The OpenAIRE Graph can support assessment-relevant analytics by enabling consistent retrieval of research products and their relations, facilitating reproducible queries, and allowing institutions to understand what is counted and why. API-mediated access is central to this functionality: it makes analytic pipelines inspectable and more readily reproducible across organisations.

6.2 Guardrails for responsible use

To remain aligned with RRA, indicator use must be explicitly bounded by guardrails: indicators should be contextualised (e.g., by discipline, mission, and career stage); limitations and uncertainty should be communicated; provenance should remain auditable to enable contestation; and the infrastructure should avoid enforcing a single evaluative model. These expectations align with both the OI4RRA checklist's ethical & inclusive foundations and the Barcelona Declaration's auditability argument.

7. Applied case: MyResearchFolio as an evaluation-facing layer on the OpenAIRE Graph

To render the infrastructural argument actionable, we conclude with MyResearchFolio as a researcher-centric service layered on the OpenAIRE Graph, offering researchers the ability to showcase their work and career advancements beyond narrow publication metrics, adding a qualitative perspective to their contributions.

MyResearchFolio enables researchers to present their research activities in a structured, interactive, and comprehensive manner. It supports both researchers and evaluators in making informed assessments related to career progression and project funding decisions. Specifically, the service provides the capability to organise research activities along a chronological timeline, explicitly linking them to research outputs, narrative contextualisation, contributions to innovation, teaching and

learning activities, and broader societal impact. In this sense, it functions as an interface that translates graph-based open research information into assessment-relevant representations while maintaining a principled separation between evidence provision and evaluative judgement. (researcherprofile.openaire.eu).

Surfacing diverse contributions in a structured yet adaptable form, and by making the underlying evidence legible rather than collapsing it into a single score, MyResearchFolio, exemplifies how open infrastructures can support plural, context-sensitive evaluation practices consistent with CoARA's reform direction and with open research information principles.

The OpenAIRE Graph represents researchers as a persistent entity that connects individuals to their scholarly outputs, affiliations, funding, and collaboration networks. These representations are constructed by reconciling trusted sources such as ORCID, institutional repositories, CRIS systems, and publication metadata, enabling accurate attribution across heterogeneous systems. Linking researchers to a broad range of outputs, including publications, datasets, and software, the OpenAIRE Graph supports assessment practices that extend beyond publication-based indicators and align with Open Science principles.

MyResearchFolio builds directly on this representation layer to deliver holistic, assessment-ready researcher profiles. Anchored in the OpenAIRE Graph, MyResearchFolio integrates interconnected metadata via the Graph's APIs and supports bidirectional synchronization with ORCID. This enables the creation of trusted, reusable profiles that can be shared across institutional and national contexts. Quantitative indicators embedded in MyResearchFolio are derived entirely from open and transparent data sources.

Crucially, MyResearchFolio augments these indicators with structured narrative CVs, allowing researchers to contextualize their contributions beyond metrics. Customizable templates support the documentation of leadership activities, interdisciplinary collaboration, mentorship, and societal engagement. By integrating qualitative narratives with open, interoperable indicators, MyResearchFolio enables libraries to implement responsible, reform-aligned research assessment practices grounded in open infrastructure, in line with the principles advocated by CoARA.

8. Conclusions and transferable takeaways for the CRIS community

This contribution argues that CRIS interoperability and responsible assessment agendas converge on a shared infrastructural requirement: open, interoperable, auditable research information. The OpenAIRE Graph can serve as a backbone layer supporting cross-institutional exchange, enrichment, and assessment-relevant evidence access, aligning with the Barcelona Declaration's call for openness and auditability and with OI4RRA WG's framework for trustworthy open infrastructures. In practical terms, the CRIS community may benefit from approaching interoperability as simultaneously technical and governance-dependent, especially regarding vocabularies and sustainability. This entails prioritising provenance and inspectability to enable contestation and trust, supporting plural and context-sensitive evaluation workflows rather than single-score benchmarking, and demonstrating value through applied layers (such as MyResearchFolio) while preserving a principled separation between infrastructure and evaluative judgement.