ADVANCES IN KNOWLEDGE ORGANIZATION
Proceedings of International ISKO Conference - ISSN: 0938-5495 - 2026 - São Paulo

Structural Integration and Semantic Limits: Persistent Identifiers in Contemporary Scholarly Communication

Raphael Figueiredo Xavier

Getulio Vargas Foundation, São Paulo State University, Brazil

Fabiano Ferreira de Castro

Federal University of São Carlos, São Paulo State University, Brazil

Abstract

Persistent identifiers have become central components of contemporary scholarly communication infrastructures and are frequently presented as technical foundations for the integration of the research ecosystem. Within this infrastructural narrative, the expansion of formal connectivity is often treated as synonymous with knowledge integration. This article critically examines that assumption from the perspective of Knowledge Organization by distinguishing between structural integration, understood as measurable and auditable formal connectivity, and semantic integration, understood as conceptual alignment and shared intelligibility. The study argues that while persistent identifiers effectively promote structural integration, they do not, by themselves, ensure semantic convergence. The analysis identifies a constitutive asymmetry between structural and conceptual integration in contemporary open science infrastructures, showing that connectivity can be scaled and monitored, whereas meaning depends on ongoing conceptual work.

1.0 Introduction

The digital transformation of scholarly communication has been understood not merely as the adoption of technological tools for publishing and disseminating research outputs, but as the emergence of a complex sociotechnical ecosystem in which digital infrastructures, technical standards, governance models, and scientific practices are inextricably interconnected (Fecher and Friesike 2014; Silva and Silveira 2019). In this context, the production, circulation, and preservation of knowledge increasingly depend on informational arrangements capable of ensuring unambiguous identification, traceability, interoperability, and long-term stability (Piwowar et al. 2018; Rosen et al. 2025).

Among the mechanisms that sustain this architecture, Persistent Identifiers (PIDs) have assumed a central role as devices for stable identification and for recording connections between scholarly entities. Initially conceived to address problems of ambiguity and persistence in the digital environment, they have gradually been incorporated into a broader infrastructural perspective in which the interlinking of records is presented as a foundation for interoperability and integration in scholarly communication (Cousijn et al. 2021).

However, the notion of interoperability mobilized in this context cannot be reduced to the technical dimension of system connectivity. While the formal connection between identifiers enables the structured circulation of information (Wilkinson et al. 2016), semantic integration depends on conceptual alignment and the construction of shared meanings (Vogt 2025). This distinction is central to the field of Knowledge Organization, which does not limit itself to connectivity between records but involves the structuring of conceptual systems.

Against this background, the following question emerges: to what extent does the expansion of formal connectivity between scholarly entities translate into semantic integration of knowledge? This article argues that, although the expansion of PIDs has produced significant advances in technical interoperability and in the formal articulation between entities, it does not in itself ensure conceptual integration.

Methodologically, this study is structured as a critical review (Grant and Booth 2009), a conceptual and theoretical approach that does not aim at exhaustive or systematic coverage of the literature, but at identifying the most significant items in a field and synthesizing them toward an original analytical distinction: in this case, between structural integration and semantic integration. The literature was therefore selected purposively and organized into three thematic clusters. The first concerns interoperability theory, used to ground the conceptual differentiation between technical/structural and semantic interoperability (Lalli 2025; Vogt 2025; Mugurusi et al. 2022; Dillen et al. 2021). The second concerns persistent identifiers as infrastructure, used to characterize the ecosystemic and integrative promise attributed to PIDs in scholarly communication (Cousijn et al. 2021; Wilkinson et al. 2016; Brown et al. 2016). The third draws on Knowledge Organization, used to anchor the epistemological argument and situate the discussion within the field's theoretical tradition (Dahlberg 1978; Svenonius 2000; Hjørland 2013, 2017; Mai 2011).

It is precisely the contrast between the ecosystemic promise identified in the second cluster and the conceptual limits discussed in the first and third clusters that evidences the asymmetry between structural and semantic integration; the distinction does not emerge from any single cluster in isolation, but from the tension between them. By proposing this analytical distinction, the study shifts the debate from a predominantly technical perspective toward the field of Knowledge Organization, highlighting that the articulation between entities does not exhaust the conceptual requirements involved in knowledge integration.

2.0 Interoperability as a Problem in Knowledge Organization

The expansion of digital infrastructures in scholarly communication has consolidated interoperability as a guiding principle of contemporary open science policies and initiatives. Understood as the ability of distinct systems to exchange and reuse information in a structured manner, interoperability is often presented as a condition for integrating data, publications, researchers, and institutions within connected ecosystems. Described as a functional capacity for structured exchange (Lalli 2025), this formulation tends to privilege

its technical dimension, associated with the standardization of formats, protocols, and identifiers, while relegating to the background the conceptual implications involved in articulating heterogeneous entities.

From the perspective of Knowledge Organization, however, the issue becomes more complex. In light of discussions on the nature of concepts and the contextuality of classification (Dahlberg 1978; Mai 2011), it becomes evident that the formal connection between records does not in itself ensure relational consistency or stability of meanings across different contexts. Knowledge integration involves processes of conceptual definition, structured representation, and interpretive mediation that go beyond the mere technical compatibility between systems (Svenonius 2000; Hjørland 2013). In this sense, to interoperate does not simply mean to exchange data, but to articulate meanings.

In this regard, interoperability may be understood as a stratified phenomenon. Recent approaches distinguish between **technical interoperability**, oriented toward structural compatibility and automated processing (Lalli 2025; Wilkinson et al. 2016), and **semantic interoperability**, related to conceptual alignment and the construction of shared intelligibility (Vogt 2025; Mugurusi et al. 2022). Structural compatibility does not necessarily correspond to convergence of meaning.

This problematization makes it possible to question the tendency to reduce interoperability to a purely technical attribute whose expansion would supposedly ensure knowledge integration. Connectivity may create conditions for integration, but it does not guarantee it.

3.0 Persistent Identifiers and the Infrastructure Imaginary of Scientific Integration

The consolidation of persistent identifiers within the ecosystem of scholarly communication did not initially occur as part of an integrated architecture of global infrastructure. Conceived to address problems of persistence and disambiguation of digital objects in distributed environments (Paskin 1999, 2005), PIDs were progressively incorporated into broader technical and institutional arrangements, eventually supporting registration, resolution, and interoperability services at an international scale (Brown et al. 2016). Through this process of institutionalization, they moved beyond a merely referential function and began to be mobilized as central components of infrastructures aimed at coordinating scientific production.

A significant portion of the recent literature shifts the analytical focus from the individual identifier to networks of identifiers conceived as the structural basis of a distributed scholarly ecosystem (Wittenburg 2019; Rosen et al. 2025). In this perspective, PIDs are presented as mechanisms that connect publications, data, researchers, organizations, and funding through persistent and machine-actionable links. The value of the identifier thus becomes associated not only with unambiguous identification but also with the capacity to structure a stable relational mesh that articulates different components of the research lifecycle.

This association becomes particularly visible in the context of the principles Findable, Accessible, Interoperable, Reusable (FAIR) and in the formulation of FAIR Digital Objects. The requirement for global, unique, and persistent identifiers as a condition for findability and interoperability (Wilkinson et al. 2016) reinforces their centrality as infrastructure for Open Science. In proposals such as FAIR Digital Objects, identifiers are embedded within the very architecture of digital objects, enabling automated discovery, access, and reuse (Schwardmann 2020).

These formulations contribute to consolidating an imaginary of sociotechnical infrastructure in which the expansion of connectivity between scholarly entities is taken as synonymous with knowledge integration. The establishment of dense networks of identifiers, associated with structured metadata and interoperable architectures, is frequently presented as the foundation of a scientific ecosystem that is articulated, transparent, and globally integrated. Integration thus appears as an expected outcome of the multiplication of formal links and the consolidation of relational structures based on PIDs.

However, this narrative tends to emphasize predominantly the structural dimension of the articulation between records and systems, without equivalently problematizing the conceptual conditions that sustain the meaning of the relationships being established. The implicit equivalence between formal connectivity and knowledge integration therefore constitutes a central assumption of this infrastructure imaginary and deserves critical examination in light of the distinction previously established between technical interoperability and semantic interoperability.

4.0 Formal Connectivity and Semantic Limits

The expansion of persistent identifiers has redefined key aspects of the informational architecture of scholarly communication. The unambiguous identification of entities, the stable resolution of references, and the possibility of establishing persistent links between distributed records have significantly expanded the technical coordination between heterogeneous systems. The consolidation of this infrastructure has enabled publications, data, researchers, organizations, and funding sources to be articulated within machine-processable relational structures, facilitating aggregation, traceability, and large-scale monitoring.

However, this structural capacity does not exhaust the issue of knowledge integration. Formal connections between entities do not, by themselves, guarantee conceptual convergence, ontological alignment, or shared meaning. In light of the distinction between technical interoperability and semantic interoperability, it becomes possible to examine more precisely both the scope and the limits of PID infrastructures, investigating to what extent the expansion of connectivity translates into effective semantic integration.

4.1 Structural Capacity of Persistent Identifiers

At the technical level, PIDs can be understood through three central properties: **unambiguous identification, persistent resolution, and structured linking between records.**

Unambiguous identification ensures that a specific entity, such as an article, dataset, or researcher, can be distinguished unequivocally in distributed digital environments. The Digital Object Identifier (DOI) assigns each object a unique identifier registered within a coordinated infrastructure, enabling consistent recognition across multiple systems (Paskin 2005). Similarly, the Open Researcher and Contributor ID (ORCID) provides persistent identifiers for researchers, reducing nominal ambiguity and enabling the consolidation of academic trajectories across different databases.

Persistent resolution constitutes the second dimension of this capacity. Through services maintained by DOI registration agencies such as Crossref and DataCite, the identifier remains stable even when the location of the resource changes. The separation between the identifier and the URL ensures continuity of access and allows different infrastructures to operate on durable references, sustaining structural interoperability between repositories, publishing platforms, and indexing systems.

The third dimension concerns the formal linking between identified entities. When a DOI is associated with structured metadata and explicit relationships, including authorship, funding, versioning, or derivation, a network of persistent connections emerges that can be processed by machines. Services that aggregate citation relationships, link publications to funding sources, or connect different versions of scholarly outputs illustrate how PID infrastructures enable traceability and systemic aggregation at scale (Brown et al. 2016; Plomp 2020).

These mechanisms produce observable effects. They support metrics, automate editorial workflows, integrate repositories and indexing services, and enable the construction of aggregated views of scientific production. The structural integration promoted by PIDs is measurable, auditable, and technically verifiable.

4.2 Semantic Limits of Identifier Infrastructures

Although identifiers ensure unambiguous identification, stable resolution, and machine-processable connections between distributed entities, they primarily perform a referential function. In the case of the DOI, its design was oriented toward persistence and disambiguation in networked environments (Paskin 1999, 2005), rather than toward incorporating conceptual descriptions of the entities it references. The distinction between identifying and describing is therefore central to understanding the limits of this infrastructure.

The linking of entities through PIDs depends on structured metadata and relational typologies defined externally to the identifier itself. In the DataCite Metadata Schema (DataCite Metadata Working Group 2024), for example, the element relatedIdentifier establishes that when it is declared, a relationType must be assigned from a controlled

vocabulary, including categories such as IsPartOf, IsReferencedBy, or IsDerivedFrom. These categories formally structure the relationship between identified entities, but their application depends on interpretation at the moment of registration and on the conceptual framing adopted by the metadata producer. The identifier technically sustains the declared connection without carrying the meaning of the relationship itself.

This distinction between identifying and describing can be read in light of Dahlberg's concept theory (1978), already mobilized in Section 2.0. The DOI performs only the referential function, pointing unambiguously to an entity, but makes no predication about it. The characteristics that would actually constitute a concept, in the Dahlbergian sense, reside entirely in the associated metadata, not in the identifier itself.

The same logic applies to relationType categories, in light of Hjørland's domain analysis (2017), already mobilized in Section 2.0. A label such as IsDerivedFrom remains syntactically identical wherever it is applied, but its semantic content depends on contextual assumptions that vary across disciplinary communities. The label sustains the formal connection between two research objects; what that connection actually means remains dependent on a situated community of practice, not on the structure of the metadata schema in isolation.

This pattern is not exclusive to the DataCite schema. Examples from other persistent identifier systems illustrate how the asymmetry between structural connectivity and semantic integration recurs across different domains of the PID ecosystem. ORCID connects researchers to their outputs and affiliations through works, employment, and education metadata, but without semantic typing of the relationships between researchers or between researchers and institutions: two authors may share a work in ORCID without any declaration of the nature of their collaboration. ROR registers institutional affiliations in a stable manner, but the relationship between a person and an organization is treated in binary terms, without modeling the nature, duration, or role of that relationship. RAiD, in turn, aggregates project entities under a research activity identifier, but without typing the individual contributions of each entity: who did what, in what capacity, and with what responsibility remains semantically opaque in the structure of the identifier. In each of these cases, the system resolves the referential problem within its specific domain without encoding the semantics of the relationships that connect these domains to one another.

The expanded architecture of the scholarly web described by Van de Sompel et al. (2014) illustrates this point. Persistent identifiers operate as essential components of technical interoperability and connectivity between distributed resources, enabling automated aggregation and the stable circulation of references. However, the existence of structural connections does not ensure conceptual alignment between the systems that produce and interpret the associated metadata.

Recent literature explicitly recognizes this gap. The CLEAR framework (Vogt 2025) emphasizes the need for conceptual clarity, rigorous entity definition, and alignment between data models as conditions for meaningful integration. Similarly, the DiSSCo Prepare project (Dillen et al. 2021) argues that persistent identifiers constitute only one

layer of interoperability, while semantic integration depends on ontological modeling, representational alignment, and conceptual governance.

PID infrastructures strengthen technical coordination between distributed systems. However, the relational intelligibility of knowledge depends on the conceptual modeling that underpins metadata and relationship typologies. Structural integration can be audited; semantic integration remains dependent on epistemological decisions and organizational practices that extend beyond the technical layer of identification.

4.3 The Structural Asymmetry of Integration

The preceding analysis reveals a constitutive asymmetry in how scientific integration is conceived within contemporary infrastructures. On the one hand, persistent identification provides verifiable technical conditions for articulating distributed entities. The adoption of PIDs can be measured, their coverage quantified, and their links computationally explored. Structural integration thus appears as an observable and scalable property.

On the other hand, conceptual integration does not submit to the same criteria of verification. Semantic alignment, relational coherence, and categorical consistency depend on epistemological decisions and modeling practices that do not automatically stabilize through the mere presence of identifiers. The intelligibility of knowledge remains conditioned by situated conceptual mediations.

This difference produces a significant asymmetry. While structural integration is facilitated by consolidated technical standards and relatively centralized governance mechanisms, conceptual integration is distributed across disciplinary domains and diverse interpretative traditions. The former tends to be normatively prioritized due to its measurability; the latter requires continuous conceptual work.

It is important to acknowledge the existence of recent initiatives that explicitly attempt to move beyond the purely structural limits of persistent identifiers. The OpenAIRE Research Graph, one of the largest open aggregations of research metadata in the world, does not merely collect and display links between identifiers. It applies inference algorithms and full-text mining across millions of open-access publications, extracting and assigning new properties and semantic relationships that were not originally declared in the metadata. This amounts to an active expansion of the meaning associated with each connection, rather than its passive registration (Aryani et al. 2020).

Similarly, the FAIR Digital Object framework proposes Data Type Registries: open, community-reviewed repositories in which scientific communities collectively define, in a hierarchical and machine-interpretable form, what each type of digital object means and which operations it supports. This institutionalizes an explicit process of semantic governance, rather than a merely technical naming convention (Schwardmann 2020). The PID Graph itself reinforces this direction: by taking identifiers, rather than the entities they represent, as the central nodes of the graph, the infrastructure is designed to make connections between identifiers richer and more easily explorable through structured

queries, enabling inferences about relationships that would not be visible from any single identifier in isolation (Cousijn et al. 2021). Taken together, these initiatives suggest that the PID infrastructure community has already recognized the insufficiency of pure formal connectivity and is actively investing in mechanisms of semantic enrichment, which could indicate that the asymmetry described in this article is being progressively overcome by the field's own technical evolution.

Even these initiatives, however, reveal the limits of the semantic enrichment project when examined in their own terms. The Scholix framework, which standardizes how publishers and data centers declare the existence of a relationship between an article and a dataset, provides a deliberately generic relationship category, "References," intended for situations in which the exact nature of that relationship cannot be determined (Burton et al. 2017). The existence of this escape category is telling: the very technical standard responsible for semantically typing connections between research entities must anticipate, in advance, the cases in which such typing is simply not feasible. The structure survives the absence of explicit semantics; it is the semantics that remains dependent on an interpretive decision the infrastructure cannot, by itself, produce.

More tellingly still, an analysis conducted by the very architects of the PID Graph, Scholix, and the OpenAIRE Research Graph acknowledges that the schemas used by these different Open Science graphs follow application-driven interpretations of research entities, which complicates their mutual integration (Aryani et al. 2020). The solution the authors themselves propose is not a single, shared semantic model across graphs, but the managed coexistence of multiple technical "dialects." The semantic enrichment of PID graphs, therefore, does not eliminate the constitutive asymmetry between connectivity and meaning. It displaces it to a new level, at which interoperability among the enrichment mechanisms themselves remains, in itself, an open conceptual question.

In this context, the expansion of connectivity begins to function as a privileged indicator of ecosystem integration, even though it does not ensure convergence of meaning. The issue does not lie in the consolidation of identification infrastructures, but in the implicit equivalence established between relational density and conceptual articulation.

Connectivity should not be conflated with intelligibility, a distinction that situates this debate within the field of Knowledge Organization, where integration implies not only connecting entities but articulating meanings.

5.0 Final Considerations

This study began from the observation that persistent identifiers occupy a central position within contemporary infrastructures of scholarly communication and are frequently mobilized as technical foundations for the integration of the research ecosystem. Based on this diagnosis, the study sought to critically examine the recurrent association between the expansion of connectivity and the integration of knowledge, shifting the debate from a perspective predominantly focused on infrastructures toward the field of Knowledge Organization.

The analytical path developed throughout the article initially involved problematizing interoperability as an issue that goes beyond technical compatibility between systems and directly affects conceptual structures and regimes of meaning. The study then examined the historical and institutional consolidation of PIDs, highlighting how these mechanisms evolved from solutions aimed at persistence and disambiguation into structural components of global architectures of scientific articulation. Finally, the distinction between structural integration and conceptual integration made it possible to examine more precisely the scope and the limits of persistent identification infrastructures.

This structural integration, however, does not automatically translate into semantic convergence. The meaning of the relationships established between identified entities depends on conceptual modeling, ontological alignment, and interpretive mediation that are not guaranteed by the technical layer of identification.

The central contribution of the study lies in making explicit a constitutive asymmetry between structural integration and conceptual integration within contemporary open science infrastructures. Recognizing this asymmetry does not imply denying the relevance of persistent identifiers, but rather situating them within their appropriate scope. The connectivity enabled by PIDs represents an important condition for the technical coordination of the scholarly ecosystem, yet the intelligibility of knowledge remains dependent on continuous conceptual work. Ultimately, integrating science does not mean merely connecting entities, but articulating meanings.

It is worth acknowledging that the literature on semantic interoperability of Knowledge Organization Systems (KOS) offers analytical instruments complementary to those mobilized in this article, including work on controlled vocabulary alignment and SKOS mappings in the Semantic Web environment (Zeng and Mayr 2019). The articulation between that literature and the analysis of PID infrastructures developed here constitutes a relevant research agenda that lies beyond the scope of this study. The present article is delimited to examining the semantic limits of persistent identifiers as mechanisms of integration in scholarly communication, without claiming to exhaust the intersections between this infrastructure and the broader field of KOS interoperability.

AI usage statement: Generative AI tools were used to assist with the translation of the manuscript into English and with linguistic revision. The authors remain fully responsible for the content, arguments, analysis, references, and final version of the manuscript.

References

Aryani, Amir, Martin Fenner, Paolo Manghi, Andrea Mannocci, and Markus Stocker. 2020. “Open Science Graphs Must Interoperate!” In ADBIS, TPDL and EDA 2020 Common Workshops and Doctoral Consortium: International Workshops: DOING, MADEISD, SKG, BBIGAP, SIMPDA, AIMinScience 2020 and Doctoral Consortium, Lyon, France, August 25–27, 2020, Proceedings, edited by Ladjel Bellatreche, Mária Bieliková, Omar Boussaïd,

Barbara Catania, Jérôme Darmont, Elena Demidova, Fabien Duchateau, et al., 195–206. Communications in Computer and Information Science 1260. Cham: Springer. https://doi.org/10.1007/978-3-030-55814-7_16.

Brown, Josh, Tom Demeranville, and Alice Meadows. 2016. “Open Access in Context: Connecting Authors, Publications and Workflows Using ORCID Identifiers.” Publications 4: 30. https://doi.org/10.3390/publications4040030.

Burton, Adrian, Hylke Koers, Paolo Manghi, Markus Stocker, Martin Fenner, Amir Aryani, Sandro La Bruzzo, Michael Diepenbroek, and Uwe Schindler. 2017. “The Scholix Framework for Interoperability in Data-Literature Information Exchange.” D-Lib Magazine 23, no. 1/2. https://doi.org/10.1045/january2017-burton.

Cousijn, Helena, Ricarda Braukmann, Martin Fenner, Christine Ferguson, René van Horik, Rachael Lammey, Alice Meadows, and Simon Lambert. 2021. “Connected Research: The Potential of the PID Graph.” Patterns 2: 100180. https://doi.org/10.1016/j.patter.2020.100180.

Dahlberg, Ingetraut. 1978. “Teoria do conceito.” Ciência da Informação 7: 101–107. https://doi.org/10.18225/ci.inf.v7i2.115.

DataCite Metadata Working Group. 2024. DataCite Metadata Schema Documentation for the Publication and Citation of Research Data and Other Research Outputs. Version 4.6. DataCite e.V. https://doi.org/10.14454/mzv1-5b55.

Dillen, Mathias, Quentin Groom, Rob Cubey, Sabine von Mering, and Alex Hardisty. 2021. A Best Practice Guide for Semantic Enhancement and Improvement of Semantic Interoperability. DiSSCo Prepare Deliverable 5.4. https://doi.org/10.34960/ajxs-zr25.

Fecher, Benedikt, and Sascha Friesike. 2014. “Open Science: One Term, Five Schools of Thought.” In Opening Science, edited by Sönke Bartling and Sascha Friesike, 17–47. Cham: Springer. https://doi.org/10.1007/978-3-319-00026-8_2.

Grant, Maria J., and Andrew Booth. 2009. “A Typology of Reviews: An Analysis of 14 Review Types and Associated Methodologies.” Health Information & Libraries Journal 26: 91–108. https://doi.org/10.1111/j.1471-1842.2009.00848.x.

Hjørland, Birger. 2013. “Theories of Knowledge Organization—Theories of Knowledge.” Knowledge Organization 40: 169–181. https://doi.org/10.5771/0943-7444-2013-3-169.

Hjørland, Birger. 2017. “Domain Analysis.” Knowledge Organization 44: 436–464. https://doi.org/10.5771/0943-7444-2017-6-436.

Lalli, Giada. 2025. “Defining Interoperability: A Universal Standard.” Journal of ICT Standardization 13: 139–156. https://doi.org/10.13052/jicts2245-800X.1323.

Mai, Jens-Erik. 2011. “The Modernity of Classification.” Journal of Documentation 67: 710–730. https://doi.org/10.1108/00220411111145061.

Mugurusi, Godfrey, Jurga Vestertė, David Asamoah, Pankaj Khatiwada, Christina Marie Mitcheltree, Halvor Holtskog, and Stian Underbekken. 2022. "The Significance and Barriers to Organizational Interoperability in Smart Service Ecosystems: A Socio-technical Systems Approach." In Advances in Production Management Systems. Smart Manufacturing and Logistics Systems: Turning Ideas into Action: IFIP WG 5.7 International Conference, APMS 2022, Gyeongju, South Korea, September 25–29, 2022, Proceedings, Part II, edited by Duck Young Kim, Gregor von Cieminski, and David Romero, 253–261. IFIP Advances in Information and Communication Technology 664. Cham: Springer. https://doi.org/10.1007/978-3-031-16411-8_31.

Paskin, Norman. 1999. "Toward Unique Identifiers." Proceedings of the IEEE 87: 1208–1227. https://doi.org/10.1109/5.771073.

Paskin, Norman. 2005. "Digital Object Identifiers for Scientific Data." Data Science Journal 4: 12–20. https://doi.org/10.2481/dsj.4.12.

Piwowar, Heather, Jason Priem, Vincent Larivière, Juan Pablo Alperin, Lisa Matthias, Bree Norlander, Ashley Farley, Jevin West, and Stefanie Haustein. 2018. "The State of OA: A Large-Scale Analysis of the Prevalence and Impact of Open Access Articles." PeerJ 6: e4375. https://doi.org/10.7717/peerj.4375.

Plomp, Esther. 2020. "Going Digital: Persistent Identifiers for Research Samples, Resources and Instruments." Data Science Journal 19: 46. https://doi.org/10.5334/dsj-2020-046.

Rosen, Isabel Abedrapo, and Ricardo Hartley Belmar. 2025. "Open Science Governance: The Role of Persistent Identifiers and Metadata Standards." OSF Preprints. https://doi.org/10.31219/osf.io/9h564_v3.

Schwardmann, Ulrich. 2020. "Digital Objects – FAIR Digital Objects: Which Services Are Required?" Data Science Journal 19: 15. https://doi.org/10.5334/dsj-2020-015.

Silva, Fabiano Couto Corrêa da, and Lúcia da Silveira. 2019. "O ecossistema da Ciência Aberta." Transinformação 31: e190001. https://doi.org/10.1590/2318-0889201931e190001.

Svenonius, Elaine. 2000. The Intellectual Foundation of Information Organization. Cambridge, MA: MIT Press.

Van de Sompel, Herbert, Robert Sanderson, Harihar Shankar, and Martin Klein. 2014. "Persistent Identifiers for Scholarly Assets and the Web: The Need for an Unambiguous Mapping." International Journal of Digital Curation 9: 331–342. https://doi.org/10.2218/ijdc.v9i1.320.

Vogt, Lars. 2025. "The CLEAR Principle: Organizing Data and Metadata Into Semantically Meaningful Types of FAIR Digital Objects to Increase Their Human Explorability and Cognitive Interoperability." Journal of Biomedical Semantics 16: 18. https://doi.org/10.1186/s13326-025-00340-7.

Wilkinson, Mark D., Michel Dumontier, IJsbrand Jan Aalbersberg, Gabrielle Appleton, Myles Axton, Arie Baak, Niklas Blomberg, et al. 2016. "The FAIR Guiding Principles for Scientific Data Management and Stewardship." Scientific Data 3: 160018. https://doi.org/10.1038/sdata.2016.18.

Wittenburg, Peter. 2019. "From Persistent Identifiers to Digital Objects to Make Data Science More Efficient." Data Intelligence 1: 6–21. https://doi.org/10.1162/dint_a_00004.

Zeng, Marcia Lei, and Philipp Mayr. 2019. "Knowledge Organization Systems (KOS) in the Semantic Web: A Multi-Dimensional Review." International Journal on Digital Libraries 20: 209–230. https://doi.org/10.1007/s00799-018-0241-2.