DCPAPERS

DCMI 2026

Proceedings of the International Conference on Dublin Core and Metadata Applications 2026

ISSN: 1939-1366

Authority Data as Context-Bearing Entities in Bibliographic Ontology Environments

Seunghui Lee [1]

[1] Chung-Ang University

Abstract

Authority data has long been valued in bibliographic systems mainly for its capacity to identify and describe entities with precision. This article reconsiders authority data as an autonomous, context-bearing entity whose significance extends well beyond this technical role. Specifically, its capacity to bear context is shown to operate across four dimensions—relational, temporal, epistemic, and domain-specific—which together constitute a distinct form of contextual meaning. Contextualization, the analysis shows, is not incidental but an intrinsic property of authority data itself, calling for a shift toward treating it as a context-bearing entity embedded within a structured knowledge network rather than as a mere access point.

1. Introduction

Library authority data has traditionally served as a mechanism of nomenclatural control, collocating name variants and regulating access points. The advent of linked data technologies has substantially reshaped this picture: URI assignment and RDF-based modeling have reconstituted authority data as a formally identifiable and richly describable entity within distributed knowledge systems. Prevailing scholarship, however, continues to evaluate authority control primarily in terms of retrieval efficiency [1], paying comparatively little attention to the broader semantic functions that authority entities perform in linked environments.

Addressing this gap is the present study's central aim. This study asks how authority data, within an ontology-based environment, comes to bear contextual meaning beyond its conventional role as an identifier, and what theoretical resources within knowledge organization research can illuminate this capacity. Drawing on the concept of boundary objects [2] and on the framing of knowledge organization as a form of social epistemology [3], this study develops a theoretical framework that reconceives authority data as an autonomous, context-bearing entity within ontology-based bibliographic environments.

2. Theoretical Background

The field-centric architecture of MARC imposed significant constraints on the expression of semantically rich inter-entity relationships in machine-readable form [4]. These structural limitations confined authority data within catalog system boundaries, poorly positioned for interoperability with external knowledge ecosystems [5]. The proliferation of Semantic Web and Linked Data technologies has driven a shift from record-centric to entity-centric library metadata practice — a movement from "records to things" [6]. IFLA LRM consolidates earlier FRBR, FRAD, and FRSAD frameworks

Student Forum

Available Aug 01, 2026

Correspondence to
Seunghui Lee
metaonto@gmail.com

into a unified entity-relationship model, formally constituting authority entities as autonomous units with attributes and relationships rather than mere access points [7]. BIBFRAME operationalizes this paradigm, expressing authority entities — Agent, Subject, Event — through URIs as identifiable and linkable units on the web.

The concept of *entityhood* refers to the property by which an information unit acquires standing as an autonomous knowledge unit — uniquely identifiable, formally describable, and semantically connectable [8]. Smith-Yoshimura [9] captures the consequence: the library community is moving from authority control to identity management, placing the entity rather than the authorized string at the center of practice. In existing scholarship, however, *entityhood* has been treated predominantly in terms of identifiability — the technical capacity to assign and resolve a URI. Sanfilippo [10] identifies this as a recurring lacuna: existing frameworks leave undertheorized the question of what an entity, once identified, can do within a knowledge system.

This reconceptualization can be situated within two established theoretical traditions in knowledge organization research. Bowker and Star's [2] concept of boundary objects describes artifacts flexible enough to be adapted to the needs of multiple communities while retaining a robust, shared identity across them. Authority entities arguably function in an analogous manner: a single authority record preserves a stable identifier while accommodating divergent interpretations and uses across cataloging, archival, and disciplinary communities. Hjørland's [3] framing of knowledge organization as a form of social epistemology further suggests that authority data should be understood not as a neutral technical artifact but as a socially and institutionally situated knowledge claim, produced through particular communities of practice and validated through recognized cataloging procedures. These two lenses inform the account of contextuality developed below, particularly its domain and epistemic dimensions.

3. Authority Data as Context-Bearing Entities

3.1. Three Constitutive Conditions of Entityhood

This study proposes that a more adequate account of *entityhood* requires three constitutive conditions — not merely one.

Identifiability: global, unique distinguishability through a persistent identifier. Satisfied in Linked Data environments by URI assignment. This is the condition existing literature has most thoroughly examined.

Describability: susceptibility to formal description through attributes and relationships. The RDF triple structure, in conjunction with BIBFRAME, Schema.org, and LRM, provides the formal machinery for this condition [7].

Context-bearing capacity: the capacity to instantiate and convey the relational, temporal, epistemic, and domain-specific contexts in which an entity is embedded. This is the semantic dimension of *entityhood* that existing frameworks have systematically

left untheorized, and the condition that most directly distinguishes authority data from other bibliographic identifiers.

These conditions stand in a hierarchical but mutually reinforcing relationship: identifiability is the ground of existence; describability is the expression of existence; context-bearing capacity is the meaning of existence within a knowledge network.

3.2. Four Dimensions of Contextuality

Relational context: authority entities derive meaning through their position within a network of relationships. For a personal name authority, affiliated corporate bodies, collaborative agents, and associated works constitute its relational context. BIBFRAME's relatedTo property, MARC 5XX fields, and LRM relationship structures formalize this dimension. Wikidata's relational density makes it a significant vehicle for expressing relational contextuality at scale [11].

Temporal context: authority entities are historically situated — birth and death dates, periods of activity, shifts in geographic nomenclature. RDF reification mechanisms and Wikidata qualifier structures provide formal support for temporally indexed assertions. Taniguchi [12] demonstrates how provenance and administrative metadata can be formally captured in RDF-based models, providing a technical foundation for this dimension.

Epistemic context: authority data is the product of institutionally situated knowledge production — specific agencies, cataloging codes, and source citations. Viewed through the lens of knowledge organization as social epistemology [3], authority data constitutes not a neutral record of fact but a socially warranted knowledge claim, produced and validated through institutionally recognized procedures of cataloging and authentication. This dimension corresponds to the Justify user task in FRAD and constitutes the distinctive epistemic depth of library authority data relative to general-purpose knowledge graph entities. The modeling of source citations, scope notes, and administrative metadata in machine-readable form makes this context formally articulable [10], [12].

Domain context: the same entity bears different contextual meaning depending on the knowledge domain or user community. In this respect, authority entities function as boundary objects [2]: a single authority record operates differently in a medical library than in a music archive [1], yet retains sufficient structural stability — through its persistent identifier and core attributes — to remain recognizable and reusable across these divergent communities of practice. Wikidata's lack of domain-specific curatorial frameworks limits its integration into library operations despite its disambiguation potential [13].

3.3. Implications

Authority data is not adequately described as an identifier. Where an identifier satisfies only identifiability, authority data satisfies all three conditions: it identifies, describes, and bears context. The move from "strings to things" [8] and from "authorities to

identifiers" [9] captures only the first step; this analysis calls for a further step — from identifiers to *context-bearing entities*.

The accumulated professional judgment embedded in authority records constitutes an epistemic resource difficult to replicate in automatically generated knowledge graph entries [6], [13]. Viewed through the lens of social epistemology [3], this judgment is not merely supplementary metadata but a socially warranted knowledge claim, embedded within institutionally recognized procedures of cataloging and authentication. Viewed through the lens of boundary objects [2], the same resource derives its value precisely from its capacity to remain stable enough for cross-domain reuse while flexible enough to support community-specific interpretation — a duality that existing evaluations of authority control, framed largely around retrieval efficiency [1], have not adequately captured. The contextualization function — present in FRAD as a user task — should accordingly be understood not as a functional outcome of authority data use but as an intrinsic property of authority data itself.

4. Conclusion

This study has argued that authority data in ontology-based bibliographic environments constitutes an autonomous context-bearing entity. First, *entityhood* requires three constitutive conditions—identifiability, describability, and context-bearing capacity. Existing frameworks have addressed the first two while leaving the third untheorized. Second, the contextuality of authority data is analyzable along four dimensions—relational, temporal, epistemic, and domain-specific—operating simultaneously within individual authority entities. This multidimensionality constitutes authority data's distinctive contribution to knowledge organization in linked environments.

By applying the concepts of social epistemology [3] and boundary objects [2], this analysis has demonstrated that the contextual value of authority data stems precisely from its dual nature as a socially warranted, institutionally validated knowledge claim, and as a structurally stable yet interpretively flexible artifact capable of cross-domain reuse.

Third, contextualization should be reconceived as an intrinsic property of authority data rather than a functional outcome of its use—with consequences for both cataloging theory and the design of ontology-based bibliographic systems. Cutter's aspiration [14] that catalogs enable users to find what they seek must, in an entity-based environment, be extended: catalogs should also enable users to understand the context in which entities exist and signify.

Acknowledgements

I would like to express my sincere gratitude to my advisor, Professor Seungmin Lee, for his invaluable guidance, thoughtful feedback, and continuous encouragement throughout this research.

References

[1] R. A. Wiederhold and G. F. Reeve, "Authority control today: Principles, practices, and trends," *Cataloging & Classification Quarterly*, vol. 59, no. 2–3, pp. 129–158, 2021, doi: 10.1080/01639374.2021.1881009.

[2] G. C. Bowker and S. L. Star, *Sorting Things Out: Classification and Its Consequences*. MIT Press, 1999.

[3] B. Hjørland, "Social epistemology," *Knowledge Organization*, vol. 51, no. 3, pp. 187–202, 2024, doi: 10.5771/0943-7444-2024-3-187.

[4] R. Tennant, "MARC must die," *Library Journal*, vol. 127, pp. 26–27, 2002.

[5] K. Coyle, "Library data in a modern context," *Library Technology Reports*, vol. 46, pp. 5–13, 2010, doi: 10.5860/ltr.46n1.

[6] C. J. Godby and K. Smith-Yoshimura, "From records to things," *Bulletin of the Association for Information Science and Technology*, vol. 43, no. 2, pp. 18–23, 2017.

[7] S. Zapounidou, M. Sfakakis, and C. Papatheodorou, "Representing and integrating bibliographic information in the Semantic Web," *Journal of Information Science*, vol. 43, no. 4, pp. 525–553, 2017, doi: 10.1177/0165551516650410.

[8] R. Fox, "From strings to things," *Digital Library Perspectives*, vol. 32, no. 1, pp. 2–6, 2016.

[9] K. Smith-Yoshimura, *From authorities to identifiers: Bridging the silos*. in OCLC Research Mini-Symposium. 2019. [Online]. Available: https://www.oclc.org/research/events/2019/120319-oclc-symposium-authorities-and-identifiers.html

[10] E. M. Sanfilippo, "Ontologies for information entities: State of the art and open challenges," *Applied Ontology*, vol. 16, no. 2, pp. 111–135, 2021, doi: 10.3233/AO-210246.

[11] C. Bianchini and L. Sardo, "Wikidata: A new perspective towards universal bibliographic control," *JLIS.it*, vol. 13, no. 1, pp. 291–311, 2022, doi: 10.4403/jlis.it-12725.

[12] S. Taniguchi, "Data provenance and administrative information in library linked data: Reviewing RDA in RDF, BIBFRAME, and Wikidata," *Cataloging & Classification Quarterly*, vol. 61, pp. 67–90, 2023, doi: 10.1080/01639374.2023.2178048.

[13] K. Tharani, "Much more than a mere technology: A systematic review of Wikidata in libraries," *The Journal of Academic Librarianship*, vol. 47, no. 2, 2021, doi: 10.1016/j.acalib.2021.102326.

[14] C. A. Cutter, "Rules for a Printed Dictionary Catalogue," U.S. Bureau of Education, 1876.