Designing the BMI Model for Interlinking Name Identifier Systems

Seunghui Lee [1,†] and Seungmin Lee [2,*,†]

[1] Chung-Ang University, 84 Heukseok-ro, Dongjak-gu, Seoul, South Korea
[2] Chung-Ang University, 84 Heukseok-ro, Dongjak-gu, Seoul, South Korea

Abstract
To uniquely identify authors of intellectual works, many institutions operate distinct name identifier systems based on their respective policies and objectives. While these systems share common elements, they differ significantly in terms of format and content, often resulting in confusion in author data management and access due to the coexistence of multiple identifiers for the same individual. This study examines the characteristics of existing name identifier systems and proposes a conceptual framework to enable their semantic interlinking.

Keywords
name identifier. ISNI, ORCID, BMI model

1. Introduction

Identification serves as the foundation for access. The presentation of access points for information resources has been a fundamental issue in information organization and metadata. Traditionally, libraries have provided access through author, title, and subject. Among these, author have been prioritized as a primary access point, traditionally used as a heading in MARC formats and increasingly emphasized in the current bibliographic environments, including Resource Description and Access (RDA) and BIBFRAME ontology.

Traditionally, author data has been managed through unit records in MARC formats. However, to enhance accessibility to information resources and efficiency in author- centered retrieval, various name identifier systems such as the International Standard Name Identifier (ISNI) have been developed. While these systems capture diverse aspects of authorship, they also introduce heterogeneity in format and content. Although name identifier systems target the same conceptual entity—authors, whether individuals or organizations—and share basic metadata elements, they often remain isolated, lacking semantic interlinking.

To address this issue, this study proposes a method for semantically interlinking heterogeneous name identifier systems while preserving the unique structures of each system. Focusing on widely adopted standards such as the International Standard Name Identifier (ISNI)

* Corresponding author.
† These authors contributed equally.
metaonto@gmail.com (Seunghui Lee); ableman@cau.ac.kr (Seungmin Lee)
0000-0002-4023-2442 (Seunghui Lee); 0000-0001-6516-8961 (Seungmin Lee)

https://doi.org/10.23106/dcmi.952576331

and the Open Researcher and Contributor ID (ORCID), this study aims to construct an upper-level conceptual framework for semantic interlinking that improve the management and utilization of author data.

2. Theoretical Background

2.1. Authority Control and Author Identifier Systems

Libraries have long maintained author authority records to uniquely identify authors and link them to bibliographic records[1]. However, MARC-based authority records, being rigid and strict, have limited in their compatibility with Web environment[2]. To overcome these limitations and enhance author identification, various name identifier systems such as ISNI and ORCID have been introduced.

Currently, name identifier systems focus primarily on verifying public identities of individuals or organizations rather than providing extensive descriptive metadata[3]. Instead, they act as bridge identifiers, connecting to external systems that provide richer contextual information[4].

2.2. Interlinking author identifier systems

Precise author identification is essential for interlinking information resources. Recent research has emphasized the need to link author data across institutions and domains[5], facilitating efficient verification and supporting access to variant names internationally.

However, heterogeneity in metadata formats and structures among systems present major challenges[6]. While institution-specific metadata is necessary, linking external author data enhances datasets and supports more accurate identification. Moving beyond treating author data as monolithic units, a more effective approach involves segmenting data at the metadata element level to enable interlinking.

3. Analysis of name identifier system data structures

ISNI, standardized as ISO 27729, uniquely identifies individuals and organizations involved in creative activities. It employs 28 core metadata elements to describe authors, their works, and their relationships. It also supports linked data publication[7][8].

ORCID identifies contributors in the research and academic domains. It provides user-managed entries across six sections, covering 30 core elements, including personal and funding information. It supports interoperability with other identifiers such as ISNI, Scopus Author ID, and Researcher ID[9].

Virtual International Authority File (VIAF) enhances global sharing of authority data and author disambiguation. It is interoperable with ISNI and offers Web-based access and browsing capabilities[10].

Although institutions independently operate their identifier systems, ongoing discussions emphasize the need for interoperability and collaboration[11]. Effective interlinking requires systematic structures that connect author information with works and related information resources.

Metadata-based structures enable linked data publication and facilitate the integration of fragmented author data for more robust interlinking. Despite operational diversity, these

https://doi.org/10.23106/dcmi.952576331

systems often share core data related to author description[12]. Semantic interlinking of this shared data can preserve the unique characteristics of each system while enhancing overall author identification functionality.

4. Name identifier data interlinking

4.1. Data fragmentation and interlinking structure

Despite differences in operational formats, name identifier systems share the common purpose of identifying authors of intellectual works. Therefore, interlinking should begin at the metadata element level. Author data, including persons and organizations, consists of metadata elements, attributes, and values. While these elements reflect the characteristics of each system, they fundamentally represent 'element-value pairs' describing bibliographic concepts.

These pairs can be categorized into four primary facets which commonly supported by various name identifier systems: Entity, Activity, Work, and Subject. These facets serve as semantic linking criteria that encompass and interrelate metadata elements used across different name identifier systems. To facilitate interlinking using these facets, individual author records are segmented into these four data units and categorized accordingly.

4.2. Facet-based interlinking framework

When categorizing the metadata elements associated with author entities, four categories, including Entity, Activity, Work, and Subject, were identified. Each element within these categories contributes to the identification of an author, yet no single metadata element alone can achieve this. Only when multiple elements are interlinked, a unique author entity can be properly identified. Therefore, the categorized elements were structured as facets to serve as the basis for semantic interlinking.

However, a structure capable of encompassing the relationships among these facets is necessary. To address this, this study applies the 'element–value pair' approach to interlink the elements within each facet category. This approach assumes that the values which are the meanings associated with an author at the bibliographic level should be semantically interlinked through relationships among metadata elements. This interlinking approach enables the construction of a layered interlinking model that supports integration at the intellectual level. Accordingly, this study proposes a conceptual model that structures the categorized elements into three levels of facets for comprehensive author identification: Bibliographic, Metadata, and Intellectual level.

4.3. Construction of BMI model

This study proposes the Bibliographic-Metadata-Intellectual (BMI) model, which structures interlinking across three levels:

In Bibliographic Level (B), author data is segmented into minimal units and linked through element-value pair mappings. Variant names are controlled using VIAF authority data, while unmatched data retains system-specific structures.

In Metadata Level (M), data linked at the bibliographic level are categorized and faceted into logical metadata groupings.

https://doi.org/10.23106/dcmi.952576331

In Intellectual Level (I), interlinked data is published as linked data, enabling verification and utilization through implemented element-value pair interlinking (see Figure 1).

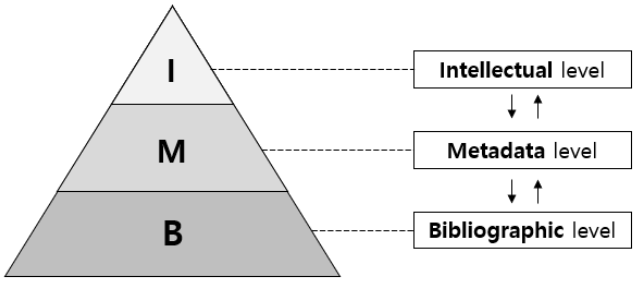

Figure 1: Basic structure of the BMI model

4.4. Interlinking structure of the BMI model

Each name identifier system is organized around the four common facets: Entity, Activity, Work, and Subject. In the BMI model, these facets are identified around an author entity, and interlinking is performed at the value level.

In this model, Bibliographic level performs element-value pair mappings. Metadata level applies categorization and facet classification. Intellectual level publishes linked data for integrated discovery and utilization.

This multi-level framework enables semantically rich interlinking while preserving the autonomy of each system. Figure 2 presents the interlinking process, emphasizing the model's goal of enhancing interoperability and usability across heterogeneous systems.

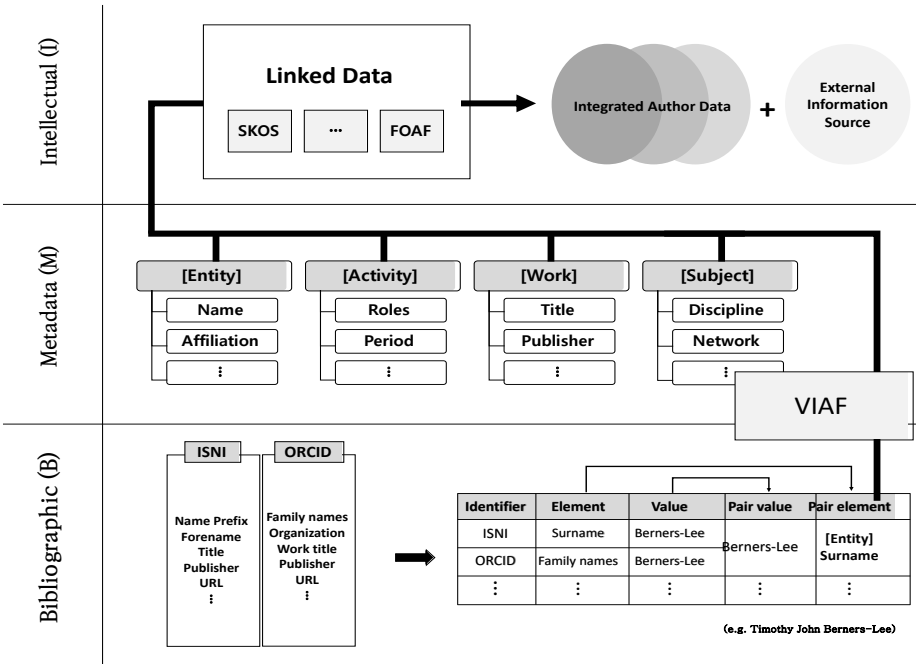

Figure 2: Interlinking structure of the BMI model

https://doi.org/10.23106/dcmi.952576331

However, each level of the BMI model and the elements included within them constitute only a conceptual structure. However, each level of the BMI model and the elements included within them constitute only a conceptual structure. For the model to be practically applied in the semantic interlinking of actual name identifiers, each of these levels needs to be implemented using a formal syntax. Accordingly, this study adopts RDF syntax to define each level and the facets as classes and properties. By defining the levels and facets of the BMI model as RDF classes and properties, the interlinking model proposed in this study can syntactically connect various name identifiers and extend them to the Web. Through this approach, metadata related to author entities can be mutually complemented, enabling richer descriptions of authors, and facilitating more accurate author identification.

5. Conclusion

This study proposes the BMI model as a semantic interlinking framework for author identifier systems that can enhance the integration and utilization of author data by interlinking shared metadata elements at the data unit level. While each system maintains its distinct characteristics, the element-value pair framework of the proposed BMI model enables categorization into Entity, Activity, Work, and Subject. facets, facilitating semantic interlinking and data sharing.

The proposed model facilitates seamless access to author information across systems, improving author data management, copyright tracking, and research evaluation. Furthermore, it supports the construction and utilization of comprehensive author authority datasets, contributing to more robust and efficient bibliographic infrastructure.

Acknowledgements

This paper is based on the author's master's thesis conducted at the Department of Library and Information Science, Chung-Ang University, South Korea.

References

[1] G. Dunsire, RDA in the fifth information age: Entity, identity, authority, Organizacija Znanja 28, no 1-2 (2023). https://doi.org/10.3359/oz2328007

[2] A. Armitage, M. J. Cuneo, I. Quintana, K. C. Young, ISNI and traditional authority work, JLIS.it 11, no 1 (2020) 151-163. doi: 10.4403/jlis.it-12554.

[3] J. Gatenby, A. MacEwan, ISNI: A New System for Name Identification, Information Standards Quarterly 23, no. 3 (2011) 4. doi: 10.3789/isqv23n3.

[4] A. MacEwan, A. Angjeli, J. Gatenby, The International Standard Name Identifier (ISNI): The Evolving Future of Name Authority Control, Cataloging & Classification Quarterly 51, no. 1-3 (2013) 55-71.

[5] L. Zhu, The future of authority control: Issues and trends in the linked data environment, Journal of Library Metadata 8 (2019). doi:10.1080/19386389.2019.1688368

[6] P. Lampron, M. Wacker, Name authority work in the linked data environment, Journal of Library Metadata 19, no. 1-2 (2019) 137-140. https://doi.org/10.1080/19386389.2019.1661109

[7] A. Angjeli, A. MacEwan, V. Boulet, ISNI and VIAF: Transforming ways of trustfully consolidating identities, in: Proceedings of the IFLA WLIC 2014, 16-22 August 2014, Lyon, France. https://library.ifla.org/id/eprint/985/1/086-angjeli-en.pdf.

https://doi.org/10.23106/dcmi.952576331

[8] V. C. Ann, S. Lieber, ISNI, a top tool for quality enhancement, smooth data flows and efficient internal processes, in: Proceedings of the IFLA WLIC 2022, 26-29 July 2022, Dublin, Ireland. https://repository.ifla.org/handle/123456789/2008.

[9] A. Dappert, A. Farquhar, R. Kotarski, K. Hewlett, Connecting the persistent identifier ecosystem: Building the technical and human infrastructure for open research, Data Science Journal 16, no 28 (2017) 1–16. https://doi.org/10.5334/dsj-2017-028.

[10] M. Downey, Assessing author identifiers: Preparing for a linked data approach to name authority control in an institutional repository context, Journal of Library Metadata 19, no. 1–2 (2019), 117–136. https://doi.org/10.1080/19386389.2019.1590936.

[11] N. Ferguson, Researcher identifiers: National approaches to ORCID and ISNI implementation, in: Proceedings of the Knowledge Exchange workshop 2015, 22-23 July 2015.

[12] S. Lee, S. Kwak, S. Oh, J. Park, A study on the management of name identifier system for ISNI-based data integration, Journal of Korean Biblia Society for Library and Information Science 30, no. 1 (2019) 29-51. doi: 10.14699/kbiblia.2019.30.1.029

https://doi.org/10.23106/dcmi.952576331