

# A BIBFRAME-Based Authority Structure for Interlinking Authority Data and Author Identifier Systems in a Linked Data Environment

Juhui Lee<sup>1</sup>   and Seungmin Lee<sup>2</sup>  

<sup>1</sup>Seoul National University, <sup>2</sup>Chung-Ang University

## Abstract

Authority data plays a critical role in clearly identifying and linking library resources, enabling users to explore more accurate and enriched information. However, current authority data remains constrained by rigid record-based structures, resulting in limited interoperability with external data sources in the web information environment. This study focuses on author authority data and proposes a method for linking it with various author identifier systems used across different domains. Within the BIBFRAME framework, which is designed for linked data environments, the technical elements of both author authority data and author identifier systems are identified at the data level. Based on this, an authority structure is conceptualized to represent diverse relationships associated with authors. The proposed structure consists of four interconnected layers centered on the author authority entity. Each component within these layers is uniquely identified by URIs and published as linked data, thereby extending traditional authority data into a web-based, interoperable ecosystem. This approach enhances the semantic connectivity of author-related information and provides a foundation for integrating heterogeneous author identification systems within the linked data environment.

## 1. Introduction

Libraries identify the authors of works and establish relationships among related resources through authority control. In a web environment where data from diverse sources coexist, extending the concept of authority control can facilitate the precise identification of author data and the explicit representation of relationships among related entities. Such an extension is essential for enabling richer information discovery and interoperability across heterogeneous data environments.

However, traditional authority data is typically constructed using complex MARC formats and organized as record-based structures. This structural characteristic limits its ability to interoperate with external data sources beyond the library domain, thereby restricting its applicability in the broader web environment.

To address this limitation, this study aims to extend traditional authority data into the web by linking it with author identifier systems utilized across various domains. Specifically, this study proposes a BIBFRAME-based authority structure that integrates the descriptive elements of author authority data and author identifier systems while explicitly representing the relationships among these elements. By employing URI-based semantic connections at the data level within a linked data environment, the proposed approach supports more effective web-based information discovery and enhances interoperability across diverse data sources.

## Full Paper

Available Aug 01, 2026

Correspondence to  
Seungmin Lee  
ableman@cau.ac.kr



Copyright © 2026 Lee & Lee. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International](https://creativecommons.org/licenses/by/4.0/) license, which enables reusers to distribute, remix, adapt, and build upon the material in any medium or format, so long as attribution is given to the creator.

## 2. Theoretical Background

### 2.1. Authority Data

Libraries have long established authority data to ensure the consistent identification and management of the resources they hold. Traditionally, authority data was primarily perceived as a tool to support cataloging activities performed by librarians. However, as users have gained direct access to authority data in online environments, it has evolved into a valuable mechanism for guiding users toward related resources [1].

In particular, authors serve as one of the most frequently used access points, functioning as a key link between bibliographic resources. Accordingly, libraries have continuously maintained author authority control to support effective search and retrieval. In this context, extending the concept of authority control to web-based author data presents an opportunity to enhance the usability and applicability of library authority data beyond the traditional library domain.

Despite its potential, authority data faces inherent limitations in expanding into external domains. Most authority data is constructed within the rigid, record-based structure of MARC formats, which restricts its interoperability with non-library data environments. Furthermore, even within the library community, authority data values are expressed in diverse forms depending on language and input conventions. The International Cataloguing Principles (2016) do not enforce a single uniform authorized form; rather, they allow the adoption of an authorized access point based on "the form most appropriate to users, including the language preferred by users or the form found in reference sources," thereby recognizing multiple authorized forms across institutions [2].

To address this diversity, the Virtual International Authority File (VIAF) has been developed as an international authority collaboration framework that links authority data from different institutions using a shared VIAF identifier [3]. By connecting variant forms that refer to the same entity, VIAF demonstrates the feasibility of integrating heterogeneous authority data across institutional and linguistic boundaries. Nevertheless, such approaches largely remain at the record level and do not fully leverage the potential of semantic, data-level integration in a linked data environment.

### 2.2. Author Identifier Systems

In recent years, various organizations involved in the management of scholarly and cultural resources have increasingly recognized the importance of author identification and have adopted author identifier systems such as ISNI and ORCID. These systems assign unique identifiers to author entities across different countries, institutions, and domains, and provide URI-based data sources that enable linkage with data from other origins [4].

Similar to how VIAF connects authority data using its own identifiers, the integration of internationally recognized author identifier systems with library authority data offers a promising approach for extending the scope of authority data. By linking authority

data with identifier systems that are widely used across domains, it becomes possible to enhance the interoperability and global connectivity of author-related information.

However, simple identifier-based linkage between authority data and author identifier systems is insufficient. Given that each system is based on different descriptive elements and structural models, a purely mechanical connection through identifiers often results in record-level linkage and may lead users to redundant or duplicated data. Such limitations hinder effective information discovery in the web environment.

To overcome these issues, a more granular and semantically rich integration is required. In a web-based information environment, authority data and author identifier systems should be connected at the level of data entities—namely, the described objects and their descriptive elements—rather than merely through identifiers. This approach enables meaningful semantic relationships among heterogeneous data sources and supports more efficient and accurate information exploration.

### *2.3. Linked Data*

The scope and descriptive practices of author data vary across domains and institutions depending on their specific needs. However, if information from heterogeneous sources can be closely interconnected on the web, users can obtain more accurate and enriched search results.

Data that is widely distributed across multiple datasets can be interconnected when published as linked data [5]. In a linked data environment, web documents are decomposed into data-level units, and each data entity is uniquely identified and connected to others, forming a large-scale semantic network. This approach enables the integration of heterogeneous data sources beyond the constraints of traditional document- or record-based structures.

In this context, both traditional authority data and author identifier systems can be reconsidered not as fixed, record-based structures but as semantically interconnected data within a linked data-friendly framework. To achieve this, each data entity must retain its own identity while being uniquely identified through URIs, allowing detailed relationships with other data entities to be explicitly represented [6]. Such an approach facilitates fine-grained semantic integration and supports more effective information discovery in the web environment.

## **3. Methodology**

### *3.1. Linking Authority Data and Author Identifier Systems in BIBFRAME*

This study adopts BIBFRAME as the framework for linking authority data with author identifier systems. BIBFRAME, introduced by the Library of Congress in 2012, is a bibliographic framework designed to maintain compatibility with MARC while transforming traditional record-based structures into data-level representations. It enables the expression of relationships among entities and supports the publication of machine-processable data [7], [8]. In BIBFRAME, data is represented as RDF graphs in which

entities are identified and interconnected through URIs, thereby facilitating integration with diverse author identifier systems during the cataloging process [4]. Furthermore, BIBFRAME has been adopted or explored as a standard bibliographic framework in several countries, including the United States, Spain, and China, indicating its growing role in international standardization and its potential to support the expansion of the authority data ecosystem.

BIBFRAME describes data through the relationships among classes and properties defined within its vocabulary. However, its structure comprises hundreds of classes and properties, making it highly complex. Simply extending the model by adding new classes or properties is insufficient to effectively support the integration of author authority data with heterogeneous author identifier systems.

To address this challenge, this study proposes a BIBFRAME-based authority structure composed of four interconnected layers. Each layer consists of classes and properties designed to link the various descriptive elements of author authority data and author identifier systems. By organizing these elements into a layered structure, the proposed model enables more systematic and semantically meaningful connections among heterogeneous data sources within the BIBFRAME framework.

### *3.2. Constructing a BIBFRAME-Based Authority Structure*

#### *3.2.1. Authority Entity*

The Authority Entity layer represents the core component of the proposed authority structure, centered on the URI of an author entity. It consists of classes and properties associated with the author entity and defines the attributes and relationships necessary for identification and linkage within the BIBFRAME framework.

In BIBFRAME, entities are identified through URIs. However, URIs themselves often lack intrinsic meaning, making them difficult for users to interpret intuitively. In web environments, users typically search for or navigate information about authors through recognizable elements such as names, identifiers, and related works. Therefore, it is necessary to associate URIs with additional descriptive components that support effective identification and linkage, while explicitly defining their attributes and relationships.

The Authority Entity layer serves as the central foundation of the overall authority structure and is positioned above the subsequent layers: Authority Entity Class, Property, and Relationship. While the Authority Entity layer focuses on the identification and linkage of author authority entities, the other three layers respectively address the structural aspects of class definition, property description, and relationship modeling. Through this layered approach, the proposed structure enables a more systematic and semantically rich representation of author-related data in a linked data environment.

### 3.2.2. Authority Entity Class

In this study, the Authority Entity Class layer is designed to function both as the target of attributes and relationships associated with the Authority Entity and as a set of nodes that enables the linkage between authority data and author identifier systems. This layer provides the conceptual categories through which author-related information is structured and semantically connected.

The components of this layer include core classes of the bibliographic domain that are related to authors, such as `bf:Agent`, `bf:Topic`, and `bf:Temporal`. These classes serve as fundamental entities for representing the contextual and semantic dimensions of author information.

However, given that this study focuses on the integration of authority data and author identifier systems for representing author information, bibliographic resources—namely Work, Instance, and Item—as well as Event entities are not included as primary components within this layer. Instead, they are incorporated as targets within the Authority Entity Relationship layer. This design choice allows the Authority Entity Class layer to remain focused on author-centric conceptual entities while delegating the representation of bibliographic and event-based connections to the relationship layer.

### 3.2.3. Authority Entity Property

The Authority Entity Property layer consists of components that represent the attributes and relationships associated with the Authority Entity Class. This layer defines the descriptive elements used to characterize author entities and to establish meaningful connections between authority data and author identifier systems.

In addition to identifiers, both authority data and author identifier systems include a wide range of descriptive elements that represent various aspects of an author. Elements that appear across both authority data and multiple identifier systems correspond to the same facets of an author and can therefore serve as key linking points between these heterogeneous data sources. Such shared elements should be incorporated into the BIBFRAME-based authority structure to enable semantic interoperability.

Furthermore, even descriptive elements that appear in only a single author identifier system may still provide valuable information for representing specific characteristics of an author. If such elements contribute meaningfully to the description of the author, they should also be included within the BIBFRAME framework.

However, even when elements refer to the same aspect of an author, their representation—such as format, granularity, and scope—may vary across authority data and different identifier systems. Therefore, it is necessary to explicitly define corresponding properties and relationships that can accommodate these variations. By doing so, the proposed structure enables a more precise and flexible semantic representation of author data within a linked data environment.

### 3.2.4. Authority Entity Relationship

The Authority Entity Relationship layer represents the relationships that an author entity can maintain, incorporating properties, classes, and relationship targets. Just as traditional library systems have linked author authority records with related bibliographic and authority records through access points, BIBFRAME requires the explicit representation of diverse relationships associated with author entities. Within BIBFRAME, relationships among bibliographic entities, related author entities, and other contextual elements enable data to be connected in a more direct and semantically rich manner.

The relationships associated with an author entity extend beyond bibliographic resources—such as `bf:Work`, `bf:Instance`, and `bf:Item`—to include relationships with other author entities (`bf:Agent`) as well as identity relationships representing the same individual across different data sources. Through these connections, the structure supports the integration of author-related information across heterogeneous datasets.

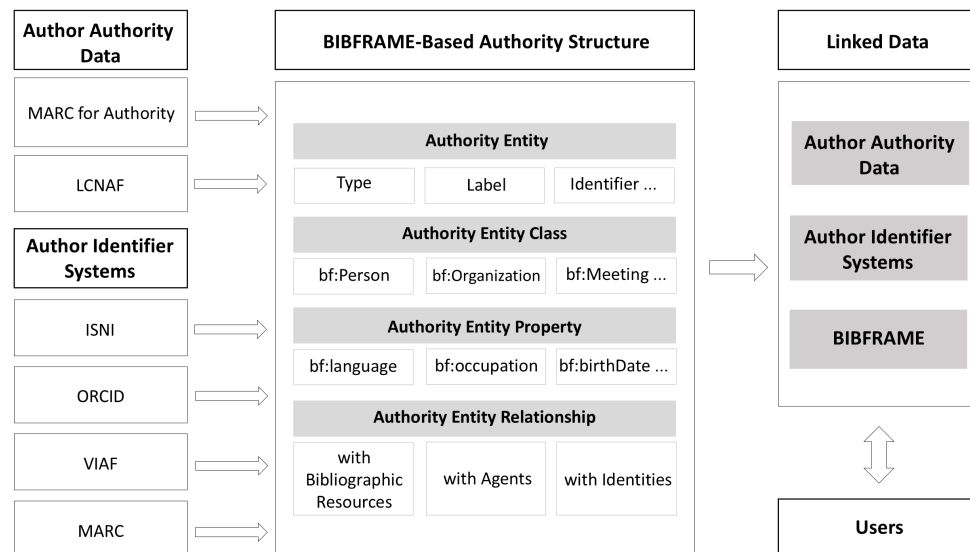
As this layer incorporates both classes and properties as its components, certain elements may overlap with those defined in the Authority Entity Class and Property layers. However, such overlaps are not redundant; rather, they reflect the multifaceted role of these elements in both defining entities and establishing relationships. By integrating these components within a relationship-oriented layer, the proposed structure enables a comprehensive and flexible representation of author-centered networks in a linked data environment.

## 4. Application of the BIBFRAME-Based Authority Structure

The integration of authority data and author identifier systems based on the BIBFRAME Authority structure proposed in this study is illustrated in [Figure 1](#). The figure presents how the four-layered structure—comprising Authority Entity, Authority Entity Class, Authority Entity Property, and Authority Entity Relationship—enables the semantic linkage of heterogeneous author-related data within a linked data environment.

As depicted in [Figure 1](#), author authority data (e.g., MARC for Authority, LCNAF) and author identifier systems (e.g., ISNI, ORCID, VIAF, and MARC-based identifiers) are currently staying in their silos, while controlling author-related resources in their own domains. To improve this situation, this study proposed a BIBFRAME-based authority structure for interlinking data from diverse sources, including authority data and author identifier systems in a linked data environment.

The structure is organized into four layers centered on URIs: Authority Entity layer, Authority Entity Class layer, Authority Entity Property layer, and Authority Entity Relationship layer. Each layer focuses on aspects of the author entity in BIBFRAME—the identification of authors, BIBFRAME Classes, BIBFRAME Properties, relationships of authors. Components within and across these layers include descriptive elements in BIBFRAME, authority data and author identifier systems. In addition to them, additional components can be newly defined for representing author-related entities.



**Figure 1.** *BIBFRAME-Based Authority Structure*

These components maintain close semantic associations and may overlap, reflecting their interconnected roles in representing author-related information. For reducing ambiguity and redundancy in the structure, attributes and relationships of each components also need to be defined in detail, based on a descriptive rule such as KCR(Korean Cataloging Rules). Through this process, these components play a key role for connecting author authority data and author identifier systems in BIBFRAME. Therefore, within this structure, authors from multiple sources can be represented, as objects of authority control in BIBFRAME.

Through this approach, the linked data is constructed as RDF graphs, based on URIs. Within the BIBFRAME framework, it is restructured and published as interoperable web data. These data can be reconnected with existing authority data and author identifier systems on the web, thereby functioning as a mediating framework that bridges authority data, identifier systems, and BIBFRAME.

Using this linked data, each institutions can complement and develop author-related data in their own domains, maintaining their natural characteristics. In addition, users can access the improved author information provided by such institutions or integrated platforms on the web. In conclusion, the authority data ecosystem is extended into the web environment, enabling users to access and explore author information from multiple sources in an integrated and semantically enriched manner.

## 5. Conclusion

To provide users with integrated access to author-related information distributed across diverse sources, this study proposed the linkage between author authority data and author identifier systems. In a linked data environment, where information traditionally organized at the document level is decomposed into data-level units and interconnected

based on the unique identification of individual data entities, authority data and author identifier systems are semantically connected.

In this context, this study designed a BIBFRAME-based authority structure that uniquely identifies author-related data through URIs of each entities. Within this structure, four layers are associated, centered on URIs. As components in each layers, data entities related with authors are included as BIBFRAME classes and properties. They represent attributes and relationships of authors, while interlinking author authority data and author identifier systems. Consequently, this structure function as a mediator which connects authority data and author identifier systems in BIBFRAME.

Through this approach, integrated data is published as linked data on the web and further interconnected with other linked datasets, ultimately forming a part of a large-scale knowledge graph. Library authority data can be extended beyond the traditional domain, linked with author identifier systems from other fields. As a result, the proposed structure enhances the interoperability of author data and supports more flexible and effective information discovery and exploration in the web environment.

This study has significance for expanding existing record-based authority data into integrated authority ecosystem that identifies and links author-related data in at the data level, within the proposed BIBFRAME-based authority structure. Through this structure, author information from various local institutions can be connected and published as linked data, generating enriched author information on the web.

## References

- [1] B. B. Tillett, "Authority control: State of the art and new perspectives," *Cataloging & Classification Quarterly*, vol. 38, no. 3–4, pp. 23–41, 2004, doi: [10.1300/J104v38n03\\_04](https://doi.org/10.1300/J104v38n03_04).
- [2] IFLA, "Statement of International Cataloguing Principles (ICP)," 2016. [Online]. Available: [https://www.ifla.org/wp-content/uploads/2019/05/assets/cataloguing/icp/icp\\_2016-en.pdf](https://www.ifla.org/wp-content/uploads/2019/05/assets/cataloguing/icp/icp_2016-en.pdf)
- [3] M. Lee, "A study on the applicability of ISNI for authority control," *Journal of the Korean Society for Information Management*, vol. 31, no. 3, pp. 133–151, 2014, doi: [10.3743/KOSIM.2014.31.3.133](https://doi.org/10.3743/KOSIM.2014.31.3.133).
- [4] J. Sandberg and Q. Jin, "How should catalogers provide authority control for journal article authors? Name identifiers in the linked data world," *Cataloging & Classification Quarterly*, vol. 54, no. 8, pp. 537–552, 2016, doi: [10.1080/01639374.2016.1238429](https://doi.org/10.1080/01639374.2016.1238429).
- [5] T. W. Cole, M. J. Han, W. F. Weathers, and E. Joyner, "Library MARC records into linked open data: Challenges and opportunities," *Journal of Library Metadata*, vol. 13, no. 2–3, pp. 163–196, 2013, doi: [10.1080/19386389.2013.826074](https://doi.org/10.1080/19386389.2013.826074).
- [6] J. Shieh and T. Reese, "The importance of identifiers in the new web environment and using the Uniform Resource Identifier (URI) in subfield zero (0): A small step that is actually a big step", *Journal of Library Metadata*, vol. 15, no. 3–4, pp. 208–226, 2015, doi: [10.1080/19386389.2015.1099981](https://doi.org/10.1080/19386389.2015.1099981).
- [7] Library of Congress, "Bibliographic framework as a web of data: Linked data model and supporting services," 2012. [Online]. Available: <https://www.loc.gov/bibframe/news/pdf/marclid-report-11-21-2012.pdf>
- [8] M. Monyela, "Library catalogue on the cloud: The BIBFRAME approach," *Journal of Library Metadata*, vol. 23, no. 1–2, pp. 59–68, 2023, doi: [10.1080/19386389.2023.2232685](https://doi.org/10.1080/19386389.2023.2232685).