Impact and Reflections of BIBFRAME Format Conversion in Foreign Libraries on China' s Cataloging Work (Postprint)

Authors: Zhang Shiyi

Date: 2025-07-09T14:55:02+00:00

Abstract

The Library of Congress announced that it will fully implement BIBFRAME cataloging beginning in 2025, standardizing catalogers' use of BIBFRAME editors for cataloging operations, which will inevitably introduce new changes to the international cataloging framework. Since most foreign language cataloging work in domestic libraries is conducted through copy cataloging of Library of Congress data, changes in its data description format will also exert a considerable impact on cataloging practices in our country. In light of this, this article examines the developmental history of BIBFRAME, the current state of research on BIBFRAME format conversion in libraries both domestically and internationally, and the challenges of format conversion. Building upon a summary of foreign experiences with BIBFRAME format conversion, the article develops reflections and ultimately proposes a series of insights for the library sector in our country to address BIBFRAME format conversion.

Full Text

Abstract

The Library of Congress (LC) has announced that it will fully implement the BIBFRAME (Bibliographic Framework) format for cataloging operations beginning in 2025, a development that will inevitably introduce significant changes to the international cataloging landscape. Since most foreign-language cataloging work in Chinese libraries relies on data sourced from LC, this shift in data format will have a non-negligible impact on domestic cataloging practices. This paper examines the developmental trajectory of BIBFRAME, reviews the current state of format conversion research both domestically and internationally, and analyzes the challenges inherent in such conversions. By summarizing lessons learned from BIBFRAME implementation experiences abroad, the study

offers reflections and recommendations for how Chinese libraries might respond to these changes, aiming to provide practical guidance for cataloging professionals navigating this transition.

Keywords: MARC; BIBFRAME; Cataloging; Format conversion

1. The Evolution of BIBFRAME

The BIBFRAME initiative has undergone substantial development since its inception. In 2012, the Library of Congress launched the BIBFRAME project, designed as a Resource Description Framework (RDF) application to replace the Machine-Readable Cataloging (MARC) format, thereby establishing a foundation for future bibliographic description. The 2015 model draft introduced three core classes, with an additional Item class added in 2021. By 2024, LC had announced its comprehensive migration plan, marking a transition from theoretical validation to practical application.

Early exploration and infrastructure development (2012–2017) saw LC collaborate with Zepheira to develop the BIBFRAME model, vocabulary, and use case documentation. This partnership produced publicly available results demonstrating modular, layered vocabulary management approaches that showcased BIBFRAME' s flexibility. The 2015 BIBFRAME pilot project represented a critical step toward transforming cataloging into linked data, enhancing library data visibility on the internet through the Libhub initiative. In 2017, the European BIBFRAME Workshop in Switzerland signaled systematic European library community engagement with linked data.

The regional cooperation and technical deepening phase (2021–2023) witnessed accelerated development of tools and platforms. OCLC' s Meridian platform enabled large-scale BIBFRAME data export and quality control, while the Linked Data for Production (LD4P) project, led by the US academic library consortium, explored decentralized cataloging models to reduce dependency on WorldCat record copying. This phase represented a shift from interoperability and data curation to emphasizing semantic expression.

The scaling and ecosystem expansion phase (2024–present) has seen BIBFRAME implementation enter a critical stage. In 2024, LC began training catalogers to input records directly in BIBFRAME format, with full FOLIO system integration planned for 2025. This timeline indicates that all LC-generated data will originate in BIBFRAME, fundamentally altering the data landscape for libraries worldwide.

2. BIBFRAME Implementation Status and International Research

2.1 International Implementation Projects

International BIBFRAME implementation has progressed through pilot projects at major institutions. The Library of Congress initiated its pilot in 2024, training 50 catalogers to create records directly in BIBFRAME format, with plans for complete FOLIO integration by 2025 and full MARC replacement by 2026. The Finnish National Library has similarly embarked on a migration plan, while the French National Library explores parallel operation of traditional and linked data cataloging workflows.

The Share Family initiative, particularly Share-VDE (Shared Virtual Discovery Environment), exemplifies collaborative approaches. This project enables BIBFRAME conversion without altering member institutions' original data, supporting multi-entity models that integrate bibliographic, authority, and holdings data. European libraries have built a cooperative network through regular workshops, addressing localization challenges such as multi-language support and regional rule compatibility.

Research on BIBFRAME' s social value demonstrates its growing adoption. Studies confirm increasing numbers of linked data projects, with BIBFRAME emerging as a promising metadata standard for replacing MARC. Scholars have analyzed differences between BIBFRAME and traditional models, highlighting BIBFRAME' s advantages in semantic expression, interoperability, and digital environment adaptation.

2.2 Domestic Research and Challenges

Domestic research has primarily focused on mapping between Chinese Machine-Readable Cataloging (CNMARC) and BIBFRAME. While no direct conversion paradigm exists, scholars have explored establishing mapping relationships, with most research remaining theoretical. Practical studies have examined field-level mappings, revealing complex relationships including one-to-one, many-to-one, and no-mapping scenarios.

Key challenges include: - Structural differences: MARC' s flat record structure versus BIBFRAME' s flexible graph-based model - Semantic mapping complexity: Control fields and indicators requiring sophisticated conversion logic - Reverse conversion difficulties: BIBFRAME' s rich semantics compressing into MARC's limited structure - Data quality risks: Potential information loss during bidirectional conversion

3. Technical Challenges in Format Conversion

3.1 MARC to BIBFRAME Conversion

The conversion from MARC to BIBFRAME presents several fundamental challenges:

Structural Incompatibility: MARC records consist of fixed-length fields, indicators, and subfields in a flat structure, while BIBFRAME uses RDF triples to express flexible entity-relationship models. This requires decomposing MARC fields into multiple BIBFRAME properties. For example, a single MARC field may need splitting across different attributes, and control codes must be transformed through complex logic.

Semantic Ambiguity: Many MARC fields lack one-to-one correspondence with BIBFRAME classes. A single field may map to multiple BIBFRAME properties, while multiple fields might consolidate into one BIBFRAME attribute. This necessitates establishing clear mapping rules during initial vocabulary alignment.

Authority Control Integration: MARC relies on text strings for authority control, whereas BIBFRAME uses URI-linked entities. Converting text entries to resolved identifiers requires matching external vocabularies (e.g., LCNAF, Wikidata), creating dependencies on external services.

3.2 BIBFRAME to MARC Reverse Conversion

Reverse conversion introduces additional complexities:

Information Compression: BIBFRAME' s rich semantic relationships may be lost when compressed into MARC' s limited structure. For instance, fine-grained creator roles expressed in BIBFRAME cannot be fully represented in MARC' s relator term fields.

Version Management: BIBFRAME' s evolving specifications create compatibility issues. Conversions must dynamically adapt to different versions, as classes and properties undergo renaming (e.g., bf:Work to bf:IntellectualProperty).

Practical Constraints: Most integrated library systems remain MARC-based, requiring BIBFRAME data to be convertible back for operational continuity. This reversible conversion reduces migration risk but demands careful handling of redundant fields and data integrity.

4. Lessons from International Implementation

4.1 Editor Development as Critical Infrastructure

International experience demonstrates that BIBFRAME editors are essential tools for catalogers. Several platforms have emerged:

- **Sinopia:** Developed by Stanford for LD4P, offering flexible vocabulary extension and multi-entity editing
- **Marva:** A next-generation editor supporting large-scale data migration and national-level collaboration
- **JCricket:** Focused on rare materials cataloging with specialized templates

These editors provide field auto-completion, authority file integration, and validation mechanisms that reduce conversion errors. For Chinese libraries, developing localized editors supporting CNMARC-specific rules (e.g., Chinese character authority control, local classification systems) is crucial.

4.2 Collaborative Implementation Models

Successful BIBFRAME adoption relies on multi-stakeholder collaboration:

Government-Industry Partnerships: LC' s collaboration with Zepheira exemplifies how clear requirements drive technical innovation. This model transformed cataloging standards into operational solutions, addressing traditional system limitations.

Regional Consortia: European libraries built cooperative networks through regular workshops, sharing localization experiences and developing shared tools like the bibfra.me vocabulary management platform.

Policy Support: National cataloging committees in countries like the UK provide policy frameworks and funding for training and tool development, creating top-down momentum for standard adoption.

4.3 Integration with RDA

The convergence of BIBFRAME (data model) and RDA (content rules) represents a core issue in modern cataloging. Finland's implementation demonstrates how extending BIBFRAME classes with RDA elements can enhance semantic expression. Regular workshops facilitate discussion of localization challenges, such as adapting multi-language cataloging rules and ensuring compatibility between international standards and local practices.

5. Strategic Reflections for Chinese Libraries

Adopting BIBFRAME in China requires addressing three interwoven challenges: technical migration, resource mismatch, and policy guidance gaps.

5.1 Overcoming System Inertia

CNMARC' s entrenched position creates path dependency. Its comprehensive coverage, standardized procedures, and deep integration into catalogers' workflows generate institutional resistance. The system continues meeting basic

service needs, reducing internal motivation for change. To address this, China should:

- Develop a national Chinese bibliographic linked data standard
- Establish a permanent authority to coordinate implementation
- Create a national exchange platform for sharing experiences
- Integrate BIBFRAME training into professional development programs

5.2 Resolving Resource Mismatches

Current training emphasizes CNMARC rules rather than linked data competencies. The gap between required semantic web skills and existing capabilities dampens adoption enthusiasm. Solutions include:

- Reconstructing cataloger competency frameworks
- Incorporating semantic technologies into library science curricula
- Developing automated tools to reduce manual conversion burden
- Piloting BIBFRAME in specialized collections (e.g., ancient books, family genealogies) where semantic enrichment offers immediate value

5.3 Policy-Driven Transformation

Top-level design must align with incremental reform. Drawing from Finland' s experience, China should:

- Issue policy statements clarifying BIBFRAME' s role in digital resource strategy
- Mandate BIBFRAME compatibility in library system procurement
- Provide dedicated funding for tool development and training
- Promote international cooperation while adapting tools to local requirements

6. Conclusion

The transition from MARC to BIBFRAME represents more than a technical upgrade—it constitutes a paradigm shift in bibliographic control. For Chinese libraries, this transformation offers opportunities to break data silos, enhance resource discoverability, and support intelligent services. However, success requires:

1. **Developing localized editors** that integrate Chinese authority files, classification systems, and cataloging rules
2. **Establishing national coordination mechanisms** to guide standardization and share best practices
3. **Implementing phased pilots** in specialized domains to build expertise and demonstrate value
4. **Aligning policy frameworks** with technical implementation to ensure sustainable development

By embracing BIBFRAME while adapting it to China' s unique bibliographic ecosystem, libraries can bridge traditional cataloging practices with semantic web technologies. This approach will not only modernize domestic cataloging operations but also position Chinese libraries as active participants in the global linked data ecosystem. The development of indigenous BIBFRAME tools and standards is both strategically necessary and practically complex, requiring vigilance against technological isolation and cost overruns while fostering an environment where library resources can achieve their full potential in the semantic web era.

Note: Figure translations are in progress. See original paper for figures.