AKR Advances on Knowledge Representation Journal

From Bibliographic Records to Linked Data: Reflections on the Transition from MARC 21 to RDA and BIBFRAME

Ana Carolina Simionato Arakaki, Felipe Augusto Arakaki

Abstract. *The transformation of representation in libraries is often described as a transition from MARC 21 to RDA and BIBFRAME. This formulation, however, brings together instruments that operate at different layers: MARC 21 is primarily an encoding and interchange format; RDA is a content standard and a set of data elements and relationships; and BIBFRAME is an RDF-based model and vocabulary. Based on a narrative literature review and a critical-conceptual synthesis, this article examines the historical development of these instruments, the shift from record-centered description to data centered on entities and relationships, and the practical conditions of migration. The analysis shows that RDA and BIBFRAME address important limitations of MARC by making entities, identifiers, and relationships more explicit and reusable. Nevertheless, the transition is neither linear nor complete. Legacy-data conversion requires interpretation and enrichment; RDA/LRM and BIBFRAME are only partially compatible. RDF may reproduce new data silos, and implementation depends on systems, governance, professional education, institutional cooperation, and sustainable identifiers. In Brazil, gradual and coordinated adoption is more feasible than the isolated replacement of formats.*

Keywords. bibliographic data; MARC 21; Resource Description and Access; BIBFRAME; linked data; interoperability.

1. Introduction

MARC 21 is one of the principal standards supporting library automation. By encoding catalog-card information in a machine-processable structure, it enabled the large-scale distribution of records in cooperative cataloging, union catalogs, and integrated library systems. Its persistence reflects stability, widespread adoption, accumulated investment, interoperability among specialized systems, and the capacity to preserve detailed descriptions across successive generations of hardware and software (McCallum, 2002).

The same stability that sustained this achievement now constrains adaptation to the Web environment. MARC 21 organizes information in records composed of tags, indicators, fields, and subfields. Although this structure can express detailed bibliographic and authority data, people, works, subjects, places, and organizations often remain embedded as textual strings within self-contained records. Their meaning depends

Adv. Know. Repr.; *2026*; volume 6; number 2. DOI: 10.5281/zenodo.22099929. Dossiê temático: O impacto das tecnologias emergentes na organização da informação e na organização do conhecimento

on combinations of field codes, punctuation, positions, cataloging rules, and local practices. Consequently, making MARC 21 data available through a website or an XML serialization does not automatically transform them into identifiable and reusable entities on the Web of Data (Alemu et al., 2012; Simionato Arakaki, 2020).

RDA and BIBFRAME emerged within this problem space, but they do not perform the same function. RDA defines elements, instructions, recording methods, and relationships for creating descriptive and access data. It can be implemented in MARC 21, relational databases, or RDF graphs. BIBFRAME, in turn, provides classes and properties for representing bibliographic data as linked resources. Treating both as equivalent "formats" that simply replace MARC obscures the layers involved in the transition and encourages an excessively linear narrative of technological succession (Bianchini & Guerrini, 2016; Silva et al., 2017).

This article argues that the transition from MARC 21 to RDA and BIBFRAME is better understood as a sociotechnical reconfiguration of bibliographic control. The most consequential change is the shift from managing monolithic records to managing entities, identifiers, relationships, vocabularies, application profiles, and services for entity identification, disambiguation, and linking. The potential benefits of this change are nevertheless constrained by three interconnected problems: the conversion and enrichment of legacy records, the partial compatibility between conceptual and implementation models, and the institutional capacity required to maintain linked data over time.

The objective is to analyze the extent to which RDA and BIBFRAME overcome structural limitations of MARC 21 and to identify the conceptual, technical, semantic, and institutional conditions that qualify this transition. The article first reconstructs the historical and conceptual foundations of cataloging. It then compares the functions and units of representation of MARC, RDA, and BIBFRAME, examines conversion and interoperability, and discusses institutional adoption, with particular attention to the Brazilian context.

2. Methodological Procedures

The study adopts a literature review combined with critical-conceptual analysis. The initial survey covered national and international publications in Portuguese, English, and Spanish, with searches in Scopus, Web of Science, and the Reference Database of Journal Articles in Information Science (BRAPCI). The time frame emphasized the period from 2008 to 2025 and included journal articles, books, doctoral dissertations, master's theses, and conference papers. The initial survey was conducted in 2023 and updated in 2026. The search strategy used terms such as metadata, cataloging, RDA, BIBFRAME, and MARC, both as acronyms and in full, in Portuguese, English, and Spanish.

Selection was guided by thematic relevance, language, recency when substantially similar versions were available, and contribution to the central problem. The materials

Adv. Know. Repr.; *2026*; volume 6; number 2. DOI: 10.5281/zenodo.22099929. Dossiê temático: O impacto das tecnologias emergentes na organização da informação e na organização do conhecimento

were read, interpreted, and organized according to seven analytical categories: (a) the nature and function of each instrument; (b) the predominant unit of representation; (c) the treatment of entities and relationships; (d) conversion and preservation of legacy data; (e) semantic and technical interoperability; (f) institutional adoption and professional competencies; and (g) evidence of value for users and services. Because the purpose is an interpretive synthesis rather than an exhaustive aggregation of evidence, the study does not claim the completeness of a systematic review.

The analysis also distinguishes claims supported by conceptual models from those supported by implementation studies. This distinction is necessary because the literature offers extensive discussion of expected benefits, such as richer navigation, reuse, and visibility, but fewer comparable assessments of operational costs and user outcomes. The resulting argument therefore treats entity-based description and linked data as significant possibilities whose realization depends on implementation choices and institutional conditions.

3. Historical and Conceptual Reconfiguration of Cataloging

Cataloging developed through a long movement from local practices toward shared principles. The rules associated with Panizzi, Jewett, and Cutter, followed by the Paris Principles and the Anglo-American Cataloguing Rules, contributed to the standardization of description and access points. AACR2, published in 1978 and articulated with the International Standard Bibliographic Description (ISBD), became an important reference for describing resources in card and automated catalogs. Even after computerization, the catalog continued to be conceived primarily as a set of descriptions whose organization reflected the logic of the catalog record (Kalita & Deka, 2020).

According to Assumpção and Santos (2015), the MARC project was developed by the Library of Congress in the 1960s in response to the practical need to encode complex cataloging data and exchange them among computers. Its leader, directory, tags, indicators, and subfields preserved the specificity of bibliographic description while supporting processing and distribution. Different national variants later emerged, creating local flexibility but also semantic and structural divergence. Convergence toward MARC 21 reduced some of this fragmentation without eliminating different cataloging policies, local fields, and uneven data quality (Kokabi, 1996; McCallum, 2002).

MARC 21, therefore, should not be reduced to an obsolete syntax. It is embedded in systems, contracts, workflows, shared-cataloging networks, and millions of legacy records. These accumulated dependencies explain why replacement is difficult and why hybrid environments are likely to persist. Any proposed successor must preserve not only fields but also decades of intellectual work, authority control, local notes, and cooperative arrangements.

Adv. Know. Repr.; *2026*; volume 6; number 2. DOI: 10.5281/zenodo.22099929. Dossiê temático: O impacto das tecnologias emergentes na organização da informação e na organização do conhecimento

3.1 FRBR, IFLA LRM, and the Centrality of Relationships

The expansion of digital resources and networked information exposed limitations that revisions of AACR2 and MARC 21 could not fully resolve. The Functional Requirements for Bibliographic Records (FRBR), published by IFLA in 1998, introduced a decisive change by modeling the bibliographic universe through Work, Expression, Manifestation, and Item. The model separated intellectual creation, realization, publication or embodiment, and the individual copy, relating these entities to the user tasks of finding, identifying, selecting, and obtaining resources (Santos & Corrêa, 2009; Kalita & Deka, 2020; Seikel & Steele, 2020).

Two further conceptual models, FRAD and FRSAD, were subsequently published, extending this entity-relationship approach to authority and subject data. Their coexistence, however, produced overlaps and inconsistencies, including different sets of user tasks and partially divergent definitions. In this context, the IFLA Library Reference Model consolidated the three models, introduced a more coherent hierarchy, and emphasized relationships as a central mechanism for navigating bibliographic information. Its user tasks—find, identify, select, obtain, and explore—reflect the movement from locating isolated descriptions to navigating networks of related resources (IFLA, 2017; Oliveira & Castro, 2022).

IFLA LRM also brought modeling in the library domain closer to terminology used in computing. Concepts such as class, superclass, subclass, domain, range, inheritance, attribute, relationship, and reification became important to the description of the model. Entities such as Res, Nomen, and Agent allow characteristics to be grouped and reused across different hierarchies. This conceptual alignment does not turn the model into an operational format, but it creates a vocabulary for designing systems in which data can be processed as explicit statements about identifiable entities (Simionato Arakaki, 2020; IFLA, 2017).

The 2016 Statement of International Cataloguing Principles reinforces this reorientation by defining users broadly enough to include technological devices that search catalogs and use bibliographic or authority data. It also introduces interoperability and openness as principles, emphasizing exchange, reuse, controlled vocabularies, automated disambiguation, and minimal restrictions on data. These principles imply that a catalog should support not only human-readable displays but also machine-processable structures (IFLA, 2016).

3.2 Semantic Web, RDF, and Data Vocabularies

The Semantic Web seeks to make the meaning of Web resources more explicit to computational agents. Linked data operationalizes this objective through practices such as identifying resources with HTTP URIs, expressing statements in RDF, and connecting local data to external datasets. RDF represents information in triples composed of subject, predicate, and object. The object may be a literal or another identified resource, allowing

Adv. Know. Repr.; *2026*; volume 6; number 2. DOI: 10.5281/zenodo.22099929. Dossiê temático: O impacto das tecnologias emergentes na organização da informação e na organização do conhecimento

a statement to become part of a graph rather than remain confined to a document or record (Bizer et al., 2009).

This structure changes the role of vocabularies. Property vocabularies define the classes and properties used to describe resources, whereas value vocabularies identify concepts, people, places, languages, genres, and other values used in statements. For linked data to support reuse and disambiguation, values should, whenever possible, be represented by persistent and resolvable identifiers rather than only by textual strings. Authority files, thesauri, classifications, code lists, and concept schemes can therefore function as shared components of bibliographic graphs (Baker et al., 2011).

The conceptual trajectory from cataloging rules to linked data can be summarized as a change in the predominant unit of management. Rules standardized descriptions; MARC transported records; FRBR and IFLA LRM modeled entities and relationships; RDA defines content at the level of elements and relationships; and BIBFRAME provides an RDF architecture for expressing these statements. Each stage responds to the limitations of the previous environment without making its predecessor immediately irrelevant.

Table 1. Historical trajectory and predominant units of bibliographic representation

Structure/standard	Primary function	Predominant unit	Limitation leading to the next stage
AACR/AACR2	Standardize description and access points	Resource description	Organization remained strongly tied to the catalog and material categories
MARC 21	Encode and exchange cataloging data	Structured record and field	Limited native integration with the Web of Data
FRBR/IFLA LRM	Model entities, relationships, and user tasks	Entity and relationship	Conceptual models alone do not provide cataloging instructions or operational encoding
RDA	Guide the creation of relational descriptive and access data	Element, entity, and relationship	Requires an encoding structure, an application profile, and suitable systems
BIBFRAME	Represent bibliographic data in RDF	Graph of identified resources	Conversion, semantic alignment, adoption, and governance remain unresolved

Taken together, the trajectory shown in Table 1 demonstrates that the transition to linked bibliographic data is not merely a change in encoding format, but a broader transformation in how library information is conceptualized, structured, and managed. The focus shifts from producing self-contained records to creating identifiable entities, explicit relationships, and reusable statements that can circulate across systems and domains. In this environment, interoperability depends not only on technical compliance with RDF but also on the consistent application of conceptual models, element sets, value vocabularies, identifiers, and governance policies. The development of linked bibliographic data should therefore be understood as a cumulative process in which

Adv. Know. Repr.; *2026*; volume 6; number 2. DOI: 10.5281/zenodo.22099929. Dossiê temático: O impacto das tecnologias emergentes na organização da informação e na organização do conhecimento

established cataloging principles are reinterpreted within a graph-based, Web-oriented infrastructure.

4. From Monolithic Records to Entities and Relationships

In MARC 21, the record is the primary transportable unit. Title, statement of responsibility, edition, publication, physical description, series, notes, subjects, and access points are brought together in a self-contained structure. Relationships are not absent: linking fields, uniform titles, added entries, notes, authority control, and identifiers may represent complex connections. The limitation lies in the fact that these relationships remain subordinate to the record and to specialized conventions for interpreting tags, indicators, subfields, positions, and punctuation.

This architecture is effective when participating systems share knowledge of MARC 21 semantics. It is less effective when external applications need to reuse a person, work, organization, or concept as an independent object. A controlled name in fields 100 or 700 may correspond to an authority record, but the bibliographic record often contains a textual form rather than a persistent identifier for the real-world entity. Consequently, the data are rich, yet relationships and identity often remain implicit.

The problem becomes more visible in legacy-data conversion. “FRBRization” requires algorithms capable of inferring works, expressions, and relationships from fields that were not originally designed to separate them unambiguously. Optional fields may be absent, uniform titles may have been applied inconsistently, notes may combine several facts, and local practices may not correspond to contemporary vocabularies. The quality of conversion is therefore determined not only by the MARC 21 specification but also by the history and regularity of each catalog (Aalberg & Žumer, 2013).

MARCXML improves syntactic accessibility but preserves the record model. XML serialization may facilitate transformation tools and APIs without automatically converting textual strings into entities or records into meaningful graphs. The relevant distinction is between publishing catalog pages on the Web and integrating bibliographic resources into the Web of Data. The latter requires explicit decisions about entities, identifiers, relationships, vocabularies, and external links.

4.1 RDA as a Content Standard and Data Vocabulary

RDA was initially published in 2010 as the successor to AACR2 and was later restructured through the RDA Toolkit Restructure and Redesign Project. Official RDA, aligned with IFLA LRM, is understood as a set of data elements, guidance, instructions, and controlled vocabularies rather than as a fixed encoding format. It supports the identification and relating of entities across different resource types and implementation scenarios (RDA Steering Committee, 2020).

Adv. Know. Repr.; *2026*; volume 6; number 2. DOI: 10.5281/zenodo.22099929. Dossiê temático: O impacto das tecnologias emergentes na organização da informação e na organização do conhecimento

RDA's most important contribution is the separation of content from a specific form of display or encoding. Descriptive decisions can be made at the level of entities, elements, and relationships, while systems determine how the data will be stored and presented. This makes it possible, at least conceptually, to reuse the same data in catalogs, repositories, discovery services, and external applications. Relationships among works, agents, translations, adaptations, editions, manifestations, and items are not mere additions to a record, but components of the description itself (Bianchini & Guerrini, 2016).

Official RDA incorporates data vocabularies and vocabularies encoding schemes. Its elements, definitions, scope notes, and controlled terms are represented in RDF and made available through the RDA Registry. The standard supports different recording methods, including unstructured descriptions, structured descriptions, identifiers, and IRIs. These alternatives enable gradual implementation but also produce different levels of machine processability. A literal name can be indexed, whereas an IRI can connect the statement to a reusable entity in a graph (RDA Steering Committee, 2020).

RDA does not require RDF or BIBFRAME. It can be applied in MARC 21, which has allowed institutions to introduce new content practices without replacing their systems. This compatibility is strategically useful, but it creates a paradox: relational data conceived through entities and relationships may be compressed into a record-oriented structure and displayed in interfaces that continue to isolate descriptions. Adopting RDA elements alone, therefore, does not necessarily constitute a transition to entity-based cataloging (Tosaka & Park, 2013).

4.2 BIBFRAME as an RDF Implementation Model

The Library of Congress announced the Bibliographic Framework Initiative in 2011 with the aim of developing a Web-adapted alternative to MARC. BIBFRAME is based on RDF, URIs, and linked-data principles. Its purpose is not simply to recode a MARC 21 record in another syntax, but to transform descriptive components into resources and relationships that can be independently identified, connected, and reused (Ramalho, 2016; Silva et al., 2017).

BIBFRAME 2.0 organizes its core around Work, Instance, and Item. Work represents the conceptual or content dimension of a resource; Instance represents its publication or material embodiment; and Item represents an individual copy. Agents, subjects, events, contributions, and other classes connect to this core through RDF properties. A contribution, for example, may be represented as a resource linking a Work to an Agent and a role, rather than being recorded as a name string in an added-entry field.

This architecture offers structural advantages: it supports persistent identifiers, treats relationships as first-class components, enables the reuse of external authorities and vocabularies, separates data management from display, and allows applications outside the library system to consume bibliographic statements. These advantages, however, are

Adv. Know. Repr.; *2026*; volume 6; number 2. DOI: 10.5281/zenodo.22099929. Dossiê temático: O impacto das tecnologias emergentes na organização da informação e na organização do conhecimento

potential rather than automatic. They depend on the effective use of resolvable URIs, documented profiles, reliable external services, and sustainable publication mechanisms.

Table 2. Functional comparison of MARC 21, RDA, and BIBFRAME 2.0

Aspect	MARC 21	RDA	BIBFRAME 2.0
Nature	Encoding and interchange format	Content standard and set of elements, guidance, and vocabularies	RDF model and vocabulary
Operational unit	Record, field, and subfield	Entity, element, and relationship	Graph of resources and properties
Identification	Control numbers, codes, and textual strings	Supports identifiers and entity identification	Favors persistent URIs
Relationships	Expressed through fields, notes, access points, and conventions	Explicitly defined as descriptive content	Expressed as RDF links
Web integration	Requires transformation and enrichment	Conceptually compatible, but implementation-dependent	Native to RDF and linked-data environments
Legacy compatibility	Very high in library systems	Can be encoded in MARC	Depends on crosswalks and conversion
Central risk	Record isolation and implicit semantics	Superficial implementation in legacy structures	Fragmented profiles and semantic loss

Table 2 highlights that MARC 21, RDA, and BIBFRAME operate at distinct yet complementary levels. MARC 21 provides a mature infrastructure for encoding and exchanging records; RDA defines the elements and relationships that constitute bibliographic description; and BIBFRAME provides a graph-based model for publishing and linking these data on the Web. The transition from MARC to BIBFRAME therefore involves more than replacing fields and subfields with RDF properties. It requires decisions about entity boundaries, identifiers, relationship semantics, vocabulary reuse, and the level of granularity adopted by an application. Because MARC records were created under different cataloging rules, local practices, and system constraints, their conversion may reveal ambiguities or implicit meanings that cannot be transferred automatically.

Consequently, the coexistence of these structures should not be interpreted only as a sequence in which one standard replaces another. In current bibliographic environments, RDA may provide descriptive content, MARC 21 may continue to support the storage and exchange of legacy data, and BIBFRAME may serve as a target model for linked-data publication. The practical relationship among them is therefore mediated by crosswalks, transformation rules, application profiles, and validation procedures. These mechanisms are essential for preserving existing data while determining which structures and semantics can be retained, enriched, reinterpreted, or lost during migration.

Adv. Know. Repr.; *2026*; volume 6; number 2. DOI: 10.5281/zenodo.22099929. Dossiê temático: O impacto das tecnologias emergentes na organização da informação e na organização do conhecimento

5. Conversion, Mapping, and Compatibility

Crosswalks establish relationships between elements in different schemas. In migration from MARC 21 to BIBFRAME, fields and subfields are associated with RDF classes and properties. This operation is indispensable for preserving legacy investment, but it cannot be understood as a neutral or lossless translation. A single MARC field may combine information belonging to several entities in the target model, while an explicit relationship in BIBFRAME may be only suggested by punctuation, order, or cataloging practice in the source record (Jett et al., 2021).

When source records do not contain persistent identifiers, automated conversion often produces literals, locally generated URIs, or blank nodes. Duplicate names may result in duplicate entities; ambiguous access points may generate incorrect links; and local notes may be reduced to generic properties. The result may be technically valid RDF that still reproduces the record-centered characteristics of MARC. Mapping should therefore document the degree of correspondence, transformation rules, inferred statements, unresolved values, and expected losses.

Practical projects demonstrate that large-scale conversion is feasible but usually requires enrichment. In a project involving approximately 300,000 e-book records, transformation to BIBFRAME was combined with queries to authority and subject services to add URIs and improve discovery. The project demonstrated technical feasibility but also revealed missing authorities, incomplete data, API failures, and maintenance requirements (Jin et al., 2016). Conversion was not merely a format change; it constituted an entity-reconciliation process dependent on external infrastructure.

The migration of the Miguel de Cervantes Virtual Library catalog followed another path: MARC 21 records were transformed into a relational model aligned with FRBR and FRAD and subsequently published in RDF using RDA vocabularies. This experience shows that BIBFRAME is not the only route to linked bibliographic data. It also reinforces the distinction between linked data as a set of principles and technologies and any specific implementation vocabulary (Romero et al., 2017).

The quality of legacy data is therefore a primary condition of migration. Auditing fields, identifying duplicates, normalizing names, reviewing authority links, and defining exception rules should precede or accompany automated transformation. Conversion projects require a combination of the competencies of catalogers, information-technology professionals, and managers: catalogers interpret historical semantics; technical teams design processing and storage; and managers define priorities, licenses, responsibilities, and sustainability.

5.1 Partial Compatibility between RDA/LRM and BIBFRAME

BIBFRAME is not a direct implementation of IFLA LRM or of RDA's entity structure. RDA and LRM distinguish Work and Expression, whereas BIBFRAME Work tends to

Adv. Know. Repr.; *2026*; volume 6; number 2. DOI: 10.5281/zenodo.22099929. Dossiê temático: O impacto das tecnologias emergentes na organização da informação e na organização do conhecimento

combine characteristics associated with both entities. BIBFRAME Instance is broadly comparable to Manifestation, and Item remains at the copy level, but even these correspondences are not complete identities (Arakaki et al., 2017). LRM entities such as Res and Nomen also have no simple equivalents in the BIBFRAME core (Seikel & Steele, 2020; Taniguchi, 2018).

The differences reflect distinct purposes. IFLA LRM is a high-level conceptual model; RDA uses that model to define content elements and relationships; and BIBFRAME was designed as an implementation vocabulary, with a strong concern for MARC-data migration. The relative simplification of Work and Expression may reduce operational obstacles because many legacy MARC records do not allow their automatic separation. At the same time, this simplification may reduce the granularity available in RDA data.

Mapping studies indicate that core entities and inherent relationships can often be represented acceptably, whereas derivative relationships—such as adaptations, translations, revisions, and transformations—may lose specificity or produce imprecise links (Zapounidou et al., 2019). The difficulty is not limited to the absence of corresponding property names. Models may define apparently similar resources according to different ontological commitments. Declaring equivalence between elements without examining their domains, ranges, and identity conditions may conceal substantial semantic differences (Hahn & Dousa, 2020; Jett et al., 2021).

Compatibility should consequently be treated as an alignment problem rather than as a one-time conversion table. Crosswalks require versioning, provenance, validation rules, and explicit statements about approximations. Institutions also need to decide which model will take priority when the same data are produced, exchanged, and displayed in multiple environments. Hybrid infrastructures may preserve MARC for exchange, apply RDA to content, and generate BIBFRAME for publication, but each additional layer creates synchronization and governance responsibilities.

5.2 Interoperability and the Risk of New Silos

RDF facilitates technical interoperability because different datasets can use a common graph model. Semantic interoperability, however, requires compatible interpretations of classes, properties, identifiers, and levels of description. Two institutions may publish valid RDF using incompatible vocabularies or local extensions that cannot be combined without new mappings. The choice between reusing established ontologies and creating collection-specific vocabularies therefore remains a persistent institutional challenge (Pennington & Cagnazzo, 2019).

The warning about moving from “MARC silos” to “linked-data silos” summarizes this problem. A BIBFRAME dataset may remain isolated when it depends on nonresolvable local URIs, uses literals where shared identifiers are available, applies undocumented profiles, or is not legally and technically accessible. The proliferation of

Adv. Know. Repr.; *2026*; volume 6; number 2. DOI: 10.5281/zenodo.22099929. Dossiê temático: O impacto das tecnologias emergentes na organização da informação e na organização do conhecimento

local ontologies may reproduce in RDF the fragmentation that the new paradigm was intended to overcome (Suominen & Hyvönen, 2017).

Identifiers are especially important because they distinguish an entity from its label. A person may have several name forms in different languages and still remain linked to the same identifier. Reusing infrastructures such as VIAF and the Library of Congress authority services may expand connections, but it also creates external dependencies, including service changes, API instability, uneven coverage, and the underrepresentation of local entities. Brazilian institutions must therefore balance international interoperability with the maintenance of identifiers and vocabularies capable of representing national authors, organizations, subjects, and cultural contexts (Castro & Jesus, 2018; Jesus & Castro, 2019).

Governance is the condition that converts technical links into sustainable connectivity. Institutions need to determine who creates and maintains URIs, how duplicate entities are merged, how changes are versioned, which licenses permit reuse, and how external applications access the data. Without shared governance, BIBFRAME may become merely another library-specific interchange format rather than an infrastructure for participation in the broader Web.

6. Institutional Capacity and Adoption

The shift toward entity-based and linked bibliographic data transforms professional work. Catalogers no longer manage only a final record; they also select or create entities, assign or reconcile identifiers, establish relationships, and apply profiles that determine mandatory and recommended properties. These activities preserve the importance of specialized cataloging knowledge while requiring additional knowledge of models, vocabularies, data quality, RDF, and automated reconciliation.

Institutional implementation requires at least five coordinated capacities. First, legacy data must be audited and improved. Second, systems must support identifiers, entity management, APIs, export, versioning, and relational or graph structures. Third, staff require continuing education that integrates cataloging principles, data modeling, and Web technologies. Fourth, governance must define licenses, persistence, provenance, responsibilities, and maintenance. Fifth, cooperation is necessary to avoid incompatible local profiles and duplicated infrastructure.

These capacities entail costs that are unevenly distributed. Large national and research libraries may test editors, converters, reconciliation services, and parallel environments, whereas smaller institutions may depend almost entirely on vendors. A transition that assumes local access to specialists, computing infrastructure, and influence over standards may increase inequalities. Cost-benefit analysis should therefore consider not only implementation expenses but also long-term maintenance, staff time, vendor dependence, and the public value generated by new services.

Adv. Know. Repr.; *2026*; volume 6; number 2. DOI: 10.5281/zenodo.22099929. Dossiê temático: O impacto das tecnologias emergentes na organização da informação e na organização do conhecimento

International adoption is heterogeneous. Cooperative initiatives involving the Library of Congress, the Program for Cooperative Cataloging, Share-VDE, and Linked Data for Production have advanced conversion practices and shared specifications, but common workflows and stable exchange arrangements are still under development (Samples & Bigelow, 2020). The existence of pilot projects and major institutional initiatives should not be interpreted as universal operational replacement.

In Spain, an institutional study identified the gradual adoption of RDA within MARC 21, while BIBFRAME remained largely experimental or prospective. This intermediate strategy allows training and policy review without the immediate replacement of systems, but it also keeps RDA's relational potential constrained by the MARC architecture (López & López-Borrull, 2018). In Canada, a survey found limited knowledge of BIBFRAME and low institutional preparedness, emphasizing the role of major libraries, professional training, and continuing education (Fortier et al., 2022).

These findings indicate that the maturity of a standard cannot be assessed solely by the stability of its specification. Operational maturity also requires commercial and open-source systems, documentation, communities of practice, exchange services, and demonstrable benefits. More recent critiques ask which next-generation workflows require BIBFRAME and which could be achieved by enriching or reorganizing MARC-based infrastructures. This does not invalidate BIBFRAME; it shifts the criterion of success from technological replacement to the provision of new and measurable capabilities (Karp, 2025).

Early studies highlighted BIBFRAME's potential to bring bibliographic data into the Web environment and recommended further research and greater professional engagement (Ramalho, 2016; Arakaki et al., 2017; Silva et al., 2017). Studies of Brazilian federal university libraries identified superficial knowledge of RDA among many professionals and the need for cooperation and training (Oliveira, 2014). More recent analyses describe national adoption as still incipient, despite pioneering institutional initiatives (Miranda & Lourenço, 2024).

The barriers combine technical, economic, linguistic, and organizational factors: integrated library systems may not fully support entities and relationships; subscription, consultancy, training, and migration costs may be prohibitive; much of the documentation is available primarily in English; and few professionals possess simultaneous expertise in cataloging, conceptual modeling, and RDF technologies. The coexistence of AACR2, RDA, and hybrid records adds another layer of complexity.

Santos and Arakaki (2022) map Latin American strategies for RDA implementation. The authors propose a set of guidelines that frame implementation as a gradual and coordinated process developed through the stages of studying, training, testing, documenting, and implementing. Initially, institutions need to study RDA's principles, structure, and underlying conceptual models and understand how it differs from previously used cataloging codes. The responsible team should then be trained

Adv. Know. Repr.; *2026*; volume 6; number 2. DOI: 10.5281/zenodo.22099929. Dossiê temático: O impacto das tecnologias emergentes na organização da informação e na organização do conhecimento

through courses, workshops, study groups, and analysis of experiences developed by other institutions. After this preparation, tests should be conducted on limited sets of records or material types to identify difficulties, adaptation needs, and effects on workflows. The results should be documented in institutional policies, manuals, record models, local guidelines, and technical procedures. Finally, implementation should proceed progressively, with the adaptation of systems, bibliographic and authority records, and cataloging routines. RDA adoption is thus a planned, collaborative, and continuous process requiring ongoing monitoring and evaluation. Several issues, however, must be considered during the transition, as shown in Table 3.

Table 3. Institutional dimensions for a gradual and sustainable transition

Dimension	Priority action	Risk addressed
Professional education	Integrate RDA, IFLA LRM, RDF, URIs, application profiles, and entity reconciliation	Nominal adoption and dependence on a small number of specialists
Legacy data	Audit fields, authorities, duplicates, notes, and local practices	Conversion that reproduces errors and ambiguous relationships
Systems	Require APIs, export, persistent identifiers, versioning, and entity-oriented functions	Vendor lock-in and the creation of new silos
Cooperation	Develop shared profiles, identifiers, services, and documentation	Fragmented practices and duplicated costs
Governance	Define licensing, URI persistence, provenance, and maintenance responsibilities	Unstable links and data without responsible curation
Evaluation	Measure discovery, reuse, productivity, data quality, and total cost	Investment driven only by technological expectations

Table 3 demonstrates that the transition to entity-based and linked bibliographic data cannot be reduced to acquiring new software or replacing one descriptive standard with another. Implementation depends on coordinated action in professional education, data management, technological infrastructure, institutional cooperation, governance, and evaluation. Weakness in any of these dimensions can compromise the entire process. A technically capable system, for example, will not produce interoperable data if identifiers are inconsistent, relationships are not recorded explicitly, or responsibilities for maintaining vocabularies and URIs remain undefined.

A gradual strategy also makes it possible to distinguish conceptual improvement from effective institutional change. Adopting RDA terminology or converting MARC records to RDF does not, by itself, ensure better discovery, reuse, interoperability, or data quality. These results depend on how standards are interpreted, how local data are reconciled, and whether institutions establish sustainable mechanisms for maintenance and cooperation. The Brazilian case therefore shows that the issue is not simply whether RDA and BIBFRAME are more advanced than MARC, but under which technical,

Adv. Know. Repr.; *2026*; volume 6; number 2. DOI: 10.5281/zenodo.22099929. Dossiê temático: O impacto das tecnologias emergentes na organização da informação e na organização do conhecimento

semantic, organizational, and economic conditions their potential advantages can be effectively realized.

7. In What Sense Do RDA and BIBFRAME Overcome the Limitations of MARC 21?

The comparison indicates a substantial but partial conceptual advance. RDA makes entities, attributes, and relationships central to descriptive practice, while BIBFRAME provides a graph architecture capable of expressing them as machine-processable links. Both weaken the assumption that the bibliographic record must be the final and indivisible unit of management. In principle, data can be recombined for different displays, services, and external applications.

The advance is partial because RDA/LRM and BIBFRAME do not model the bibliographic universe in the same way. The Work–Expression distinction, the treatment of Nomen, and the representation of derivative relationships require approximations. The new environment therefore resolves some limitations of MARC while creating a contemporary layer of alignment among models, vocabularies, profiles, and implementations.

From a technical perspective, BIBFRAME is more directly suited to RDF and linked data than MARC. Even so, the outcome depends on the effective use of URIs, external links, and stable services. A graph populated mainly by literals and local blank nodes may offer little interoperability beyond that obtained from a transformed record. MARC data can also be exposed through APIs and enriched with identifiers, which means that not every service innovation requires immediate replacement. The relevant question is whether keeping MARC 21 at the core increases costs, constrains relationships, or prevents substantially simpler workflows in a graph-native architecture.

Legacy-data conversion receives an intermediate assessment. It is technically scalable, but it involves interpretation, inference, enrichment, and review. The value of decades of cataloging cannot be preserved through syntactic conversion alone. Libraries need to decide which distinctions can be inferred, which relationships should remain uncertain, and which local data deserve representation in the target model.

Interoperability remains uncertain. RDF and shared vocabularies expand potential connectivity, while divergent profiles, local extensions, unstable identifiers, and licensing restrictions may shift rather than eliminate silos. The main determinant is not only the data model but also the governance arrangements that sustain shared interpretation and maintenance.

Operational adoption is low to intermediate because most environments combine MARC, RDA, experimental linked-data services, and vendor-specific workflows. Coexistence is not merely a temporary failure; it may constitute a rational strategy for preserving cooperative exchange while new capabilities are developed. Hybrid

Adv. Know. Repr.; *2026*; volume 6; number 2. DOI: 10.5281/zenodo.22099929. Dossiê temático: O impacto das tecnologias emergentes na organização da informação e na organização do conhecimento

infrastructures, however, must make synchronization, authority control, and the source of truth for entities and relationships explicit.

Evidence of user benefits remains less developed than the conceptual and technical literature. Explicit relationships can support the grouping of editions, translations, adaptations, and related works. In this context, linked data may extend visibility beyond the catalog, and identifiers may reduce ambiguity. Institutions nevertheless need indicators such as search precision and coverage, external reuse, reduction of duplicated work, cataloging time, maintenance cost, and the accessibility of resulting interfaces. Without evaluation, a library may replace an established infrastructure with a more complex one without demonstrating proportional public value. Table 4 synthesizes the discussions and relationships among RDA, BIBFRAME, and MARC.

Table 4. Assessment of the extent to which RDA and BIBFRAME overcome MARC limitations

Dimension	Degree of progress	Critical assessment
Conceptual model	High, but partial	Entities and relationships become central, but RDA/LRM and BIBFRAME diverge
Web suitability	High in potential	RDF and URIs support connections only when used effectively and sustainably
Legacy conversion	Intermediate	Feasible at scale, but dependent on inference, reconciliation, and enrichment
Interoperability	Uncertain	Shared technologies may improve exchange or merely relocate silos
Operational adoption	Low to intermediate	Coexistence, pilot projects, and gradual transitions predominate
Demonstrated user benefit	Still limited	Promising prototypes exist, but broad and comparable evaluations are scarce
Feasibility for smaller institutions	Problematic	Costs, competencies, and technological dependence may increase inequalities

The central point is that the transition does not consist of replacing fields with triples. It changes what libraries manage. The object of management expands from the final record to entities, identifiers, relationships, controlled vocabularies, application profiles, provenance, reconciliation processes, and data services. This new environment is sociotechnical: without bibliographic analysis, automation may create incorrect links at scale; without automation, the volume and connectivity expected of linked data are not viable.

Consequently, MARC 21, RDA, and BIBFRAME should be analyzed as layered and potentially coexisting instruments. RDA may improve the semantic quality of data while they remain encoded in MARC 21; BIBFRAME may publish or exchange graph-based data derived from different content standards; and MARC 21 may remain part of

Adv. Know. Repr.; *2026*; volume 6; number 2. DOI: 10.5281/zenodo.22099929. Dossiê temático: O impacto das tecnologias emergentes na organização da informação e na organização do conhecimento

cooperative workflows during a prolonged transition. The strategic objective should not be the disappearance of MARC 21 as an end in itself, but the development of bibliographic capabilities that preserve intellectual investment while enabling more explicit, reusable, and connected data.

8. Final Considerations

RDA and BIBFRAME address important limitations of MARC 21, but they do not replace it through a single, linear operation. RDA provides a more flexible and relational basis for descriptive content by separating elements and relationships from a specific encoding. BIBFRAME provides an RDF architecture in which bibliographic resources can be identified, linked, and reused as a graph. Together, they support a conception of catalogs as networks of entities rather than merely as collections of records.

The most significant advance is conceptual. The unit of management shifts from the monolithic record to entities, identifiers, and relationships. Operational transition is considerably more difficult. Legacy data contain implicit semantics and local variation, while crosswalks may involve approximation and consequent data loss. Official RDA and BIBFRAME are not fully equivalent, and RDF may reproduce new silos when identifiers, profiles, and governance are fragmented.

Institutional capacity, therefore, is not an external detail of implementation but part of the model's practical feasibility. Sustainable adoption requires data auditing, professional education, appropriate systems, persistent identifiers, documented profiles, licensing, cooperation, and evaluation of outcomes for users and organizations. These requirements are particularly important for smaller institutions and for countries in which systems, training resources, and national coordination are uneven.

The studies discussed tend to favor gradual and cooperative strategies because they are more feasible than isolated full conversion. Importing a model without building governance may simply replace MARC 21 silos with RDF silos. The transition should therefore be used to review data, relationships, and services while strengthening social reuse and the Web presence of library collections.

The decisive question, therefore, is not when MARC 21 will disappear, but which bibliographic capabilities need to be preserved, transformed, or created. RDA and BIBFRAME are relevant instruments for this process when treated as components of a sociotechnical infrastructure rather than as self-sufficient technological solutions.

References

Aalberg, T., & Žumer, M. (2013). The value of MARC data, or, challenges of frbrisation. Journal of Documentation, 69(6), 851–872. https://doi.org/10.1108/JD-05-2012-0053

Alemu, G., Stevens, B., Ross, P., & Chandler, J. (2012). Linked data for libraries: Benefits of a conceptual shift from library-specific record structures to RDF-based data models. New Library World, 113(11/12), 549–570. https://doi.org/10.1108/03074801211282920

Arakaki, F. A., Galeffi, L. F., Alves, R. C. V., & Santos, P. L. V. A. da C. (2017). BIBFRAME: tendência para a representação bibliográfica na web. Revista Brasileira De Biblioteconomia E Documentação, 13, 2231–2249. Retrieved from https://rbbd.febab.org.br/rbbd/article/view/995

Assumpção, F. S., & Santos, P. L. V. A. C. Representação no domínio bibliográfico: um olhar sobre os Formatos MARC 21. (2015). Perspectivas Em Ciência Da Informação, 20(1), 54-74. https://periodicos.ufmg.br/index.php/pci/article/view/22971

Baker, T., Coyle, K., & Petiya, S. (2011). Library Linked Data Incubator Group final report. World Wide Web Consortium. http://www.w3.org/2005/Incubator/lld/XGR-lld-20111025/

Bianchini, C., & Guerrini, M. (2016). RDA: A content standard to ensure the quality of data. JLIS.it, 7(2), 83–98. https://doi.org/10.4403/JLIS.IT-11709

Bizer, C., Heath, T., & Berners-Lee, T. (2009). Linked Data: The story so far. International Journal on Semantic Web and Information Systems, 5(3), 1–22.

Castro, F. F. de, & Jesus, A. F. de. (2018). Adequando dados bibliográficos ao Linked Data: Requisitos necessários, vantagens e desafios. Informação & Sociedade: Estudos, 28(3). https://doi.org/10.22478/ufpb.1809-4783.2018v28n3.38288

Fortier, A., Pretty, H. J., & Scott, D. B. (2022). Assessing the readiness for and knowledge of BIBFRAME in Canadian libraries. Cataloging & Classification Quarterly, 60, 708–735. https://doi.org/10.1080/01639374.2022.2119456

Hahn, J., & Dousa, T. (2020). Mapping bf:Work to lrm:Work and lrm:Expression: Towards a set-theoretical approach. Proceedings of the Association for Information Science and Technology, 57. https://doi.org/10.1002/pra2.408

International Federation of Library Associations and Institutions. (2016). Statement of international cataloguing principles. https://www.ifla.org/files/assets/cataloguing/icp/icp_2016-en.pdf

International Federation of Library Associations and Institutions. (2017). IFLA Library Reference Model: A conceptual model for bibliographic information. https://www.ifla.org/files/assets/cataloguing/frbr-lrm/iflalrm-august-2017_rev201712.pdf

Jesus, A. F. de, & Castro, F. F. de. (2019). Dados bibliográficos para o linked data. Brazilian Journal of Information Science: Research Trends, 13(1), 45–58. https://doi.org/10.36311/1981-1640.2019.V13N1.08.P45

Jett, J., Dubin, D., & Bothmann, B. (2021). To map or not to map: Rethinking crosswalk agendas. NASKO, 8(1). https://doi.org/10.7152/NASKO.V8I1.15889

Jin, Q., Hahn, J. F., & Croll, G. (2016). BIBFRAME transformation for enhanced discovery. Library Resources & Technical Services, 60(4), 223–235. https://doi.org/10.5860/LRTS.60N4.223

Kalita, D., & Deka, D. (2020). Searching the great metadata timeline: A review of library metadata standards from linear cataloguing rules to ontology-inspired metadata standards. Library Hi Tech, 39(1), 190–204. https://doi.org/10.1108/LHT-08-2019-0168

Karp, S. (2025). BIBFRAME without BIBFRAME: The infrastructure of next-generation cataloging workflows. Technical Services Quarterly, 42, 318–327. https://doi.org/10.1080/07317131.2025.2512281

Kokabi, M. (1996). The internationalization of MARC. Part I: The emergence and divergence of MARC. OCLC Systems & Services, 12(1), 21–31. https://doi.org/10.1108/10650759610118622

López, R. C., & López-Borrull, A. (2018). Estudio de la adaptación a RDA y BIBFRAME en el ámbito de las bibliotecas españolas. Anales de Documentación, 21(2). https://doi.org/10.6018/ANALESDOC.21.2.323441

McCallum, S. (2002). MARC: Keystone for library automation. IEEE Annals of the History of Computing, 24(2), 34–49. https://doi.org/10.1109/MAHC.2002.1010068

Miranda, D. R. S., & Lourenço, C. de A. (2024). RDA: Recomendações de implementação no Brasil. Biblios Journal of Librarianship and Information Science. https://doi.org/10.5195/biblios.2024.1185

Oliveira, M. A. L. de. (2014). RDA: Utilização do novo código de catalogação pelas bibliotecas das universidades federais brasileiras [Dissertation].

Oliveira, R. H. A. de, & Castro, F. F. de. (2022). Contribuições do IFLA LRM para o RDA: Uma revisão sistemática da literatura. Brazilian Journal of Information Science: Research Trends, 16, e02156. https://doi.org/10.36311/1981-1640.2022.v16.e02156

Pennington, D., & Cagnazzo, L. (2019). Connecting the silos: Implementations and perceptions of linked data across European libraries. Journal of Documentation, 75(3), 643–666. https://doi.org/10.1108/JD-07-2018-0117

Ramalho, R. A. S. (2016). BIBFRAME: Modelo de dados interligados para bibliotecas. Informação & Informação, 21(2), 292–306. https://doi.org/10.5433/1981-8920.2016V21N2P292

RDA Steering Committee. (2020). RDA Toolkit. https://www.rdatoolkit.org/

Romero, G. C., Esteban, M. P. E., Carrasco, R. C., & Such, M. (2017). Migration of a library catalogue into RDA linked open data. Semantic Web, 9, 481–491. https://doi.org/10.3233/SW-170274

Samples, J., & Bigelow, I. (2020). MARC to BIBFRAME: Converting the PCC to linked data. Cataloging & Classification Quarterly, 58, 403–417. https://doi.org/10.1080/01639374.2020.1751764

Santos, A. da S. dos, & Arakaki, F. A. (2022). Aplicação do RDA em instituições latino-americanas: um panorama sobre o processo de implementação. Revista Brasileira De Biblioteconomia E Documentação, 18, 1–18. Retrieved from https://rbbd.febab.org.br/rbbd/article/view/1678

Santos, P.L.V.A.C., & Corrêa, R.M.R. Catalogação: trajetória para um código internacional. Intertexto, 2009.

Seikel, M., & Steele, T. (2020). Comparison of key entities within bibliographic conceptual models and implementations: Definitions, evolution, and relationships. Library Resources & Technical Services, 64(2), 62–71. https://doi.org/10.5860/LRTS.64N2.62

Silva, L. C. da, Segundo, J. E. S., Zafalon, Z. R., & Santos, P. L. V. A. da C. (2017). O código RDA e a iniciativa BIBFRAME: Tendências da representação da informação no domínio bibliográfico. Em Questão, 23(3), 130–156. https://doi.org/10.19132/1808-5245233.130-156

Simionato Arakaki, A. C. (2020). O modelo IFLA Library Reference Model e o Linked Data. Informação & Informação, 25(3), 163–186. https://doi.org/10.5433/1981-8920.2020v25n3p163

Suominen, O., & Hyvönen, N. (2017). From MARC silos to linked data silos. O-Bib. Das Offene Bibliotheksjournal, 4(2), 1–13. https://doi.org/10.5282/O-BIB/2017H2S1-13

Taniguchi, S. (2018). Mapping and merging of IFLA Library Reference Model and BIBFRAME 2.0. Cataloging & Classification Quarterly, 56, 427–454. https://doi.org/10.1080/01639374.2018.1501457

Tosaka, Y., & Park, J. (2013). RDA: Resource Description & Access—A survey of the current state of the art. Journal of the Association for Information Science and Technology, 64, 651–662. https://doi.org/10.1002/asi.22825

Zapounidou, S., Sfakakis, M., & Papatheodorou, C. (2019). Mapping derivative relationships from RDA to BIBFRAME 2. Cataloging & Classification Quarterly, 57, 278–308. https://doi.org/10.1080/01639374.2019.1650152