

Article

# IFLA Library Reference Model and Schema.org for Semantic Enrichment of Library Catalogs

Dayane Onaga Ferreira Machado<sup>1,\*,\dagger</sup>, Ana Carolina Simionato Arakaki<sup>1,2,\*,\dagger</sup><sup>1</sup>Department of Information Science, Federal University of São Carlos (UFSCar), São Carlos, SP 13565-905, Brazil<sup>2</sup>Faculty of Information Science, University of Brasília (UnB), Brasília, DF 70910-900, Brazil\*Correspondence: [onagadayane@gmail.com](mailto:onagadayane@gmail.com) (Dayane Onaga Ferreira Machado); [ana.arakaki@unb.br](mailto:ana.arakaki@unb.br) (Ana Carolina Simionato Arakaki)<sup>†</sup>These authors contributed equally.

Academic Editor: Natália Tognoli

Submitted: 11 October 2025 Revised: 8 December 2025 Accepted: 31 January 2026 Published: 24 April 2026



Dayane Onaga Ferreira Machado is a Ph.D. student in Information Science at the Federal University of São Carlos (UFSCar), Brazil. She is currently a researcher in a project funded by the Federal District Research Foundation (FAP-DF). Her research focuses on the Semantic Web, conceptual models, ontologies, and Linked Data.



Ana Carolina Simionato Arakaki is an Associate Professor at the Faculty of Information Science, University of Brasília (UnB). She is a member of the graduate programs in Information Science at the University of Brasília and at the Federal University of São Carlos (UFSCar) and is a CNPq Research Productivity Fellow (Level 2) in Brazil. Her research focuses on knowledge organization, metadata, linked open data, conceptual models, digital curation, and data management. She coordinates funded research projects dedicated to interoperability and the qualified use of authority and bibliographic data in digital environments and supervises undergraduate and graduate students in Information Science.

## Abstract

Libraries are widely recognized for providing relevant and reliable information through digital catalogs. However, the traditional formats used in cataloging do not support effective data sharing on the web. To address current informational demands and expand access, the modernization of catalogs is essential. This article proposes a theoretical alignment between the International Federation of Library Associations and Institutions Library Reference Model (IFLA LRM) and Schema.org. Specifically, it establishes a crosswalk mapping attributes from IFLA LRM entities (Work, Expression, Manifestation, and Item) to Schema.org's CreativeWork properties, aiming to support the semantic enrichment of library catalogs. The research adopts a qualitative, exploratory approach, using the crosswalk technique to map attributes of IFLA LRM entities (Work, Expression, Manifestation, and Item) to Schema.org CreativeWork properties. The Library Reference Model provides a strong foundation for bibliographic data modeling, enabling alignment with metadata standards and semantic vocabularies. However, the entities and attributes of IFLA LRM are not fully compatible with the types and properties defined in Schema.org, as the models serve different purposes. Even so, both IFLA LRM and Schema.org offer flexible data structures that allow for adaptation. This study addresses a gap in the literature on interoperability between IFLA LRM and Schema.org and offers insights into their potential alignment for the semantic enrichment of library catalogs. Aligning these models may promote greater openness and interoperability with web search engines, thereby enhancing user experience and increasing interaction with and exploration of library catalogs.

**Keywords:** library catalog; digital catalog; library reference model; Schema.org; semantic enrichment; crosswalk

## 1. Introduction

Digital catalogs play a crucial role in libraries by enabling access to and retrieval of reliable and relevant information resources. With technological advancements, new informational demands have emerged, requiring the continuous updating of these instruments. In this context, catalog modernization can be achieved through system modeling

grounded in conceptual frameworks specifically designed for the bibliographic domain.

The International Federation of Library Associations and Institutions (IFLA) Library Reference Model (LRM) is a high-level conceptual model developed to guide the formulation of cataloging rules and the implementation of bibliographic systems. Its primary objective is to reduce ambi-



guity and to allow the assignment of namespaces for Linked Open Data (LOD) applications (Riva et al., 2017). As noted by Arakaki (2020), through the instantiation of IFLA LRM and the creation of namespaces, it becomes possible to establish mappings with other metadata standards.

Beyond data linkage, it is equally important to consider the discoverability of catalog content on the web. From this perspective, aligning catalog modeling with semantic vocabularies such as Schema.org can support data enrichment and integration on the web. Developed by the major search engines, Schema.org enhances data interpretation, generating more precise and relevant search results and thereby improving user experience.

It has been observed that, although many studies have addressed different aspects of catalog modernization, a significant gap remains in the literature concerning the association between IFLA LRM and Schema.org. By examining the convergence between IFLA LRM and Schema.org, this article not only contributes to the modernization of cataloging practices but also responds to broader challenges of interoperability, data openness, and user-centered access in the era of the Semantic Web. Accordingly, through an exploratory and descriptive research approach, this article examines bibliographic data modeling from the perspective of IFLA LRM, as well as the potential alignment of this model with Schema.org.

This research is situated at the intersection of bibliographic standards and semantic web technologies, addressing the pressing need for interoperable solutions that enhance the visibility and usability of library data in a global digital ecosystem.

## 2. Methodology

This research adopts a qualitative, exploratory approach with a theoretical focus. This study is framed as exploratory conceptual modeling research. Establishing a theoretical framework is a prerequisite for empirical validation, as it helps identify the variables and relationships to be tested. Consequently, this article focuses on the conceptual coherence and feasibility of the alignment, providing the groundwork for future prototype development and user experience studies. For data analysis, the crosswalk technique was used, as originally proposed by the National Information Standards Organization (NISO) in 1999 (Ballinger, 2015). The procedure involved a comparative analysis of conceptual models in the bibliographic domain. The attributes of the IFLA LRM, specifically for the entities Work, Expression, Manifestation, and Item, were examined and mapped to the properties of the Schema.org CreativeWork type. Equivalences were established primarily by comparing the official definitions from the IFLA LRM specification and the Schema.org documentation (version 29.3). Additionally, the examples in the IFLA LRM specification served as complementary evidence to validate the scope and practical interpretation of each attribute, ensuring closer alignment with the intended function of the cor-

responding Schema.org properties. The analysis was carried out manually through documentary examination of the models.

## 3. Conceptual Modeling Results From the IFLA Library Reference Model Perspective

Considering technological developments, rethinking the instruments that govern the descriptive treatment of information resources is crucial for building more efficient systems aligned with the needs of new generations of users. The representation of information resources encompasses the selected descriptive elements, retrieval tools, and record structures, promoting “[...] portability, reliability, management, consistency, accuracy, and relevance of the results obtained in response to a query” (Zafalon, 2017, p. 10).

Conceptual models guide the organization and representation of information within a domain. By establishing a robust conceptual framework, these models allow for a reexamination of resource description, making explicit the main entities, the attributes that characterize them, and the ways in which these elements interrelate.

To establish logical connections within a database, the conceptual structure may be based on the Entity–Relationship Model (ERM), originating from Computer Science. This method enables the description of entities, attributes, and relationships. The cardinality of relationships is a key element of ERM, as it defines the minimum and maximum number of connections between instances of entities, typically represented as one-to-many (1:M), many-to-many (M:M), or one-to-one (1:1).

In the bibliographic context, conceptual models shift the focus from static records to a data-centric view, allowing information to be represented through entities and relationships. From this standpoint, the IFLA Library Reference Model (LRM) emerges as a set of guidelines for bibliographic domain representation. The LRM was designed to harmonize the Functional Requirements (FR) models, which include Functional Requirements for Bibliographic Records (FRBR), focused on bibliographic records; Functional Requirements for Authority Data (FRAD), specialized in authority records; and Functional Requirements for Subject Authority Data (FRSAD), centered on subject records.

Conceptual models are formulated from the users’ perspective, considering the tasks they perform in a system and how they retrieve the desired information. As explained by Riva et al. (2017), the functional scope of the IFLA LRM is based on the main tasks performed by end and intermediary users: to find, identify, select, obtain, and explore. The task find refers to meeting search criteria; identify concerns distinguishing between similar resources; select refers to evaluating the suitability of retrieved resources according to user preferences; obtain addresses access to the resource; and finally, explore relates to the discovery of resources through context and interrelationships.

**Table 1. International Federation of Library Associations and Institutions (IFLA) Library Reference Model entities.**

Entity	Entity hierarchy	Definition
Library Reference Model Entity (LRM E1): Res	Top level	Any entity in the universe of discourse.
LRM E2: Work	Second level	The intellectual or artistic content of a distinct creation.
LRM E3: Expression	Second level	A distinct combination of signs conveying intellectual or artistic content.
LRM E4: Manifestation	Second level	A set of all carriers that are assumed to share the same characteristics as to intellectual or artistic content and aspects of physical form.
LRM E5: Item	Second level	An object or objects carrying signs intended to convey intellectual or artistic content.
LRM E6: Agent	Second level	An entity capable of deliberate actions, of being granted rights, and of being held accountable for its actions.
LRM E7: Person	Third Level	An individual human being.
LRM E8: Collective Agent	Third Level	An entity capable of deliberate actions, of being granted rights, and of being held accountable for its actions.
LRM E9: Nomen	Second level	An association between an entity and a designation that refers to it.
LRM E10: Place	Second level	A given extent of space.
LRM E11: Time-span	Second level	A temporal extent having a beginning, an end and a duration.

The IFLA LRM provides a high-level theoretical framework. However, as stated by Riva et al. (2017), it is not a cataloging code nor an encoding schema. Therefore, its operationalization requires the mediation of implementation rules (such as Resource Description and Access (RDA)) and serialization formats (such as Resource Description Framework (RDF)) to become machine-actionable in web environments. This distinction creates a necessary translation layer between the abstract model and the concrete data description. Table 1 presents the IFLA LRM entities, their hierarchies, and corresponding definitions.

Table 1 presents all the entities of the model. An entity is characterized as an abstract class of conceptual objects. All entities are accompanied by a definition. The top-level entity is termed Res (thing), with the following direct subclasses: Work, Expression, Manifestation, Item, Agent, Person, Collective Agent, Nomen, Place, and Time-span. Although hierarchically equal to other second-level entities, the core structural entities of the model are Work, Expression, Manifestation, and Item (Riva et al., 2017). Fig. 1 illustrates the relationships and cardinalities among them.

Fig. 1 above illustrates the relationships among the entities, with cardinality indicated by the arrowheads: a single arrowhead denotes a one-to-one (1) relationship, while a double arrowhead indicates a many (M) relationship. Thus, it can be understood that a Work is realized through one or more Expressions; an Expression is the realization of only one Work; an Expression may be embodied in more than one Manifestation, which in turn may be exemplified by one or more Items, each represented by a single Manifestation. In other words, a Work is the intellectual or artistic creation conceived by an individual or an organization, and the Expression is the realization of that work, which can be conveyed in different formats and languages; the Manifestation materialized in an Item, such as, for instance, a book.

The conceptual structure of the IFLA LRM differs from the approach of traditional standards such as Machine-Readable Cataloging (MARC) 21 and cataloging rules such as Anglo-American Cataloguing Rules (AACR2r). While these methods rely on rigid, flat record structures, the LRM organizes bibliographic data through interconnected entities and relationships.

The IFLA LRM was designed for the context of the Semantic Web, whose purpose is to provide context to data, enabling efficient interpretation by computational systems. In this way, modeling aligned with the IFLA LRM makes it possible to implement more efficient bibliographic systems, reducing ambiguity in representation and allowing for the assignment of namespaces for Linked Open Data (LOD) applications. Linked Open Data consists of a set of principles for creating meaningful information networks through the publication of machine-readable structured data linked to other data (European Union, 2022).

The Statement of International Cataloguing Principles (ICP) addresses general principles of cataloguing, among which Principle 2.10 highlights the interoperability of catalog data, both within and beyond the library context. According to the Statement, “All efforts should be made to ensure the sharing and reuse of bibliographic and authority data within and outside the library community. For the exchange of data and discovery tools, the use of vocabularies facilitating automatic translation and disambiguation is highly recommended” (IFLA, 2016, p. 5).

Therefore, domain modeling based on the definitions and structure of the IFLA LRM enables the application of Linked Data principles. In addition to linking data, it is essential to ensure the discoverability of catalog records on the web, as described in Cataloguing Principle 2.10. To achieve this, data must be standardized according to structured data vocabularies and semantic markup standards rec-

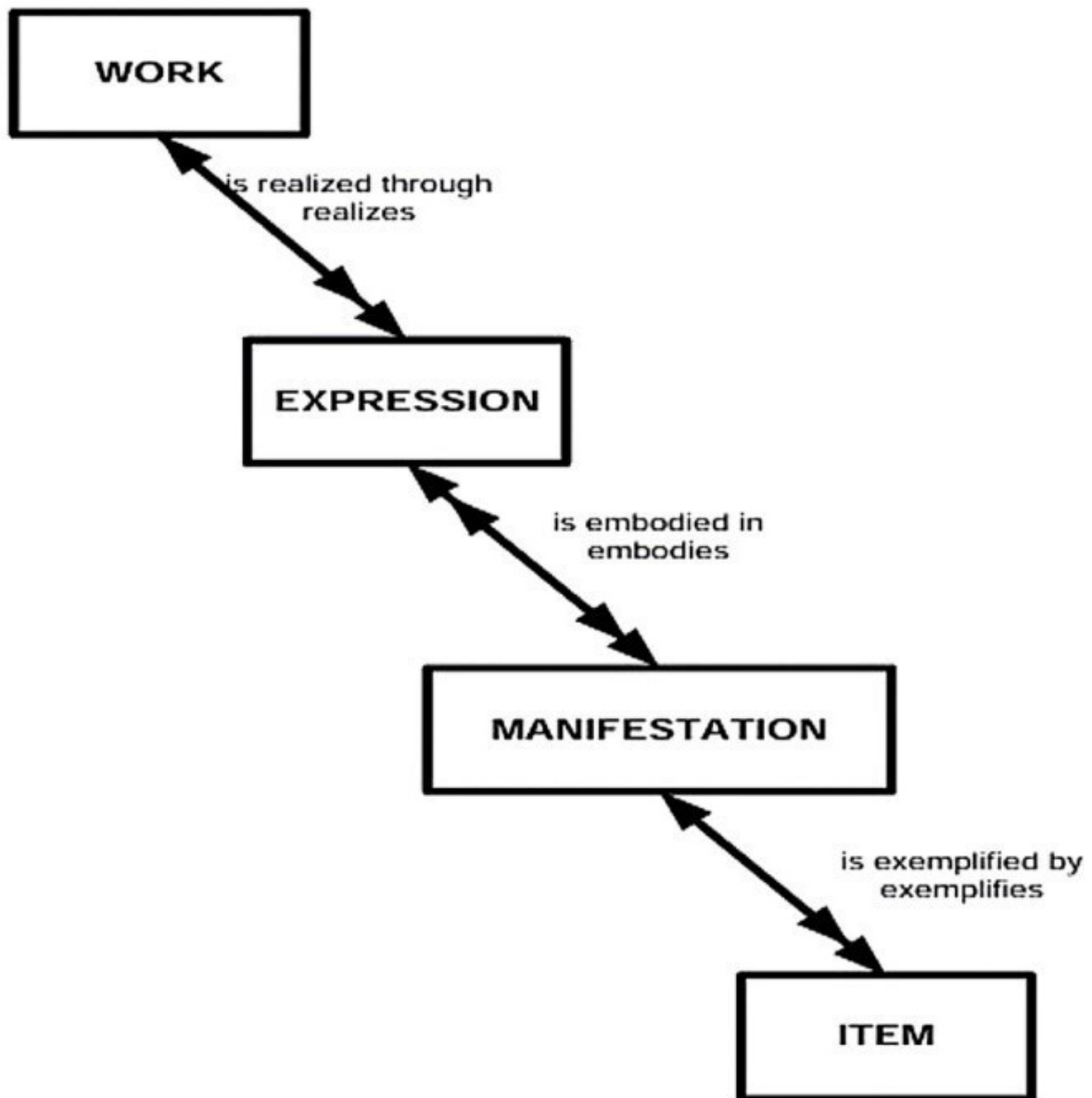


Fig. 1. Relationships between work, expression, manifestation, and item.

ognized by major search engines. From this perspective, the next section addresses the Schema.org semantic vocabulary.

#### 4. Schema.org

Schema.org is a semantic vocabulary developed by Google, Microsoft, Yahoo, and Yandex with the purpose of creating, maintaining, and promoting schemas for structured data (Schema.org, 2025a). Schema.org provides a set of schemas designed to standardize the structuring of web pages, thereby facilitating their discovery and interpretation by search engines. The vocabulary is managed collaboratively, involving both representatives of the founding companies and the broader web community.

In addition to its benefits for information retrieval, Schema.org emphasizes the importance of using a common

vocabulary to standardize and harmonize data description on the web. This approach simplifies the implementation process for developers: “A shared vocabulary makes it easier for webmasters and developers to decide on a schema and get the maximum benefit for their efforts” (Schema.org, 2025a).

The vocabulary consists of a set of schemas, referred to as types or classes, each associated with a set of properties. These elements are organized within a multiple inheritance hierarchy, allowing each type to function as a subclass of multiple types simultaneously. Thing is the top-level class, encompassing the following subclasses: Action, BioChemEntity, CreativeWork, Event, Intangible, MedicalEntity, Organization, Person, Place, Product, and Taxon.

In addition to the class hierarchy, there is also the DataTypes hierarchy, which defines the types of values

that can be assigned to class properties. This hierarchy includes: boolean, covering the values false and true, useful for expressing binary conditions or states; date, which includes dates in ISO 8601 format (year-month-day); date-time, representing the combination of date and time; number, referring to numerical values and including specific types such as float and integer; text, which represents textual values (including CSSSelectorType, Pronounceable-Text, URL, and XPathType); and time, used to express specific times of day.

The data model employed by Schema.org is compatible with widely adopted formats and semantic encodings for structured data, such as Resource Description Framework in Attributes (RDFa), Microdata, and JavaScript Object Notation for Linked Data (JSON-LD). By integrating Schema.org into web pages using these recommended formats, search results may be incorporated into the Knowledge Graph. This database system collects, organizes, and establishes relationships among information about entities such as people, places, and things, thereby enhancing search performance and delivering more precise and relevant results to users (Google, 2024).

At the time this article was written, Schema.org was in version 29.3. According to its official website, it is estimated that more than 45 million web pages had already adopted Schema.org (Schema.org, 2025a). This widespread adoption can be attributed to the vocabulary's generic and adaptable nature, which highlights its potential for application in diverse contexts. Such adaptability is evidenced by its use across multiple platforms and contexts: Pinterest applies tags to implement Rich Pins (Pinterest, 2023); Facebook enriches shared links (Meta, 2025); Cortana (Windows 10 and Windows Phone) identifies tags in email messages (Guha et al., 2016); Apple leverages it in search functionalities through Siri (Guha et al., 2016); and Amazon employs structured data to interpret and respond to user queries through Alexa (Amazon, 2023).

As noted by Iliadis et al. (2025), Schema.org represents a unique opportunity to examine the emergence of a global metadata vocabulary, one that may play an important role in how web data are shared and displayed across major applications and platforms.

Despite its broad adoption, Schema.org has been criticized for prioritizing commercial and general web contexts, which may limit its ability to fully capture the complexity of bibliographic and cultural heritage data. Beyond these applications, there are initiatives implementing Schema.org within the library domain to enhance the semantic enrichment of bibliographic records, such as those led by the Online Computer Library Center (OCLC) and the National Library Board Singapore. The following section discusses these initiatives and highlights the benefits achieved. The discussion should be presented separately from the results and focus on interpreting the findings rather than restating

them. Avoid excessive speculation and ensure that relevant literature is cited to support the interpretation. Toward the end of the discussion, briefly address the strengths and weaknesses of the research.

### *Schema.org in the Bibliographic Domain*

OCLC was a pioneer in exploring the use of Schema.org within the library context. The vocabulary was incorporated into WorldCat's data structuring projects. According to Fons et al. (2012), OCLC was already engaged in Linked Data initiatives at the time Schema.org was launched, which allowed for the enhancement of these projects with a focus on exposing structured data on the web. The OCLC initiatives demonstrate that aligning bibliographic data with Schema.org is feasible and impactful. However, they also highlight the importance of extending the vocabulary to accommodate the complexity of bibliographic records, reinforcing the role of libraries as data providers in the semantic web ecosystem.

In this context, OCLC developed an experimental model for structuring WorldCat data based on Linked Data and Schema.org. This initiative paralleled the development of Bibliographic Framework (BIBFRAME), the Library of Congress's linked data successor to MARC 21, which grounds bibliographic data in RDF and FRBR-family models (Godby, 2013). Although BIBFRAME modernizes data exchange within the library community, OCLC's work demonstrates that alignment with Schema.org remains essential for achieving broad visibility and interoperability on the general web.

The stages and outcomes of these projects were described by Godby (2013); Godby and Denenberg (2015); Godby et al. (2015); and Godby (2016). According to Godby (2016), developers considered that the CreativeWork class could be structured into RDF triples, thereby establishing a semantic model for representing the main informational resources of libraries. This class includes relevant metadata for describing a wide range of resources such as books, articles, academic works, paintings, sculptures, manuals, photographs, musical recordings, films, series, games, among others.

However, its generic orientation raises questions about the adequacy of Schema.org for domains requiring high levels of precision and contextual richness, such as libraries, archives, and museums.

In partnership with the World Wide Web Consortium (W3C), OCLC created a working group to extend Schema.org, named the Schema Bib Extend Community Group, which focused on promoting greater granularity for bibliographic data (World Wide Web Consortium, 2025).

Schema Bib Extend developed BiblioGraph.net (BGN), an extension vocabulary with additional properties for the CreativeWork class. These properties were subsequently incorporated into the official vocabulary, and the terms created by this group are currently identified in the official schema with the extension bib.schema.org

([Bibliograph.net, 2019](#); [Schema.org, 2025b](#)). Godby (2016) highlights that this extension allowed libraries to include concepts that are intelligible both to the cataloging community and to other domains.

The vocabulary was designed based on entities defined by OCLC, and its relationships were structured according to the FRBR conceptual model, focused on Group 1 entities: Work, Expression, Manifestation, and Item. It is important to emphasize that the concept of Work in Schema.org differs from that defined in the FRBR and LRM models. Among the main entities mentioned, Work stands out for its distinct definition in the Schema.org context. Godby (2013, p. 11) notes that, within libraries, a Work is understood as: “[...] much more abstract—it is a unique intellectual endeavor with an ascribed authorship, which is imperfectly represented in library systems as a cluster of descriptions, or a set of properties that are common among the editions, formats, or translations.” In contrast, Schema.org defines work from the perspective of web users, treating it as a tangible object, such as a book or DVD (Godby, 2013).

From this perspective, according to Schema.org (2025c), “[...] something might simultaneously be both a Book and a Product and be usefully described with properties from both types. It is useful but not required for the relevant types to be included in such a description.” Thus, the relationships between classes, subclasses, and properties in Schema.org are not rigid rules but rather recommendations on how the vocabulary may be reused and combined depending on context.

A practical example of this application can be observed in OCLC’s use of Schema.org in WorldCat. By incorporating structured data markup into its catalog pages, WorldCat enables search engines such as Google to interpret and index its bibliographic records. For instance, a Google search for *Rules for a Dictionary Catalog* by Charles Ammi Cutter retrieves a WorldCat record directly within the search results, as illustrated in Fig. 2 (OCLC, 2024a). This demonstrates how the adoption of Schema.org contributes to the web visibility of library catalog records.

Fig. 2 shows a WorldCat bibliographic record retrieved on the first page of search results. In the source code of this record, Schema.org markup in JSON-LD format was identified, as illustrated in Fig. 3 (OCLC, 2024b).

The Schema.org markup shown in Fig. 3 is organized within a structure called DataFeed. A DataFeed is a single feed that provides structured information about one or more entities or topics. The property dataFeedElement represents an item within the data feed in this case, the book being searched. The markup indicates the correspondence between the book and the Book entity, and between the author and the Person entity. In addition, the properties define information such as the book title (name), the URL, the format (bookFormat), the language (inLanguage), and the International Standard Book Number (ISBN).

The data structuring illustrated in Fig. 3 enables the integration of this content into the Knowledge Graph. Such

structuring supports the retrieval of information as rich snippets, which highlight the most relevant details directly on the search results page. Fig. 4 presents an example of an enriched result obtained from a search for *Alice’s Adventures in Wonderland* by Lewis Carroll.

In Fig. 4, the search result for the book *Alice’s Adventures in Wonderland* is displayed directly on the results page, showing options such as overview, summary, characters, get book, analysis, reviews, more by author, and quotes. By selecting “get book” (highlighted in red at the top of the figure), a field is displayed (highlighted in red at the bottom of the figure) where the user can enter their postal code to check whether the item is available in a nearby library.

A relevant observation is that when the city name “São Paulo” is entered in the search field, a message appears (the last highlight in red in the figure) indicating that no results were found for the requested resource. However, this message suggests updating the postal code or conducting the search directly in WorldCat. Therefore, even if the initial search does not return results from a catalog, the user is directed to WorldCat, enabling the exploration of collection contents and the discovery of other resources of interest.

To make these functionalities available for libraries, Google provides specific instructions for page coding. To activate the lending option, it is recommended to create separate feeds: one for the book entity and another for the library entity. These feeds must employ the recommended formats for structured data and the properties defined by Schema.org (Google Search Central, 2025).

Regarding the application of Schema.org in the bibliographic domain, during the LD4 Conference on Linked Data in 2023, Richard Wallis, one of the main contributors to the Schema.org initiative, presented the data structuring project of the National Library Board Singapore (NLB). The author discussed the application of linked data together with BIBFRAME for the NLB, a government organization in Singapore responsible for promoting reading, information literacy, and interest in the nation’s history and heritage (National Library Board, 2025).

According to Wallis (2023), Schema.org was crucial to the data model developed in this project. Data in MARC-XML, DC-XML, and CSV formats were aligned with BIBFRAME to preserve the integrity and detail of bibliographic records, while Schema.org was employed to describe the entities, thus facilitating discovery and interpretation by search engines. The results, according to the author, demonstrated the successful transformation and integration of NLB data into a structured format suitable for analysis and exploration within knowledge graphs (Wallis, 2023). Beyond institutional implementations, academic and working group initiatives have further explored these alignments. The LIBER Linked Open Data Working Group (2021) established best practices for semantic interoperability, highlighting the use of vocabularies like Simple Knowledge Organization System (SKOS) for concept description

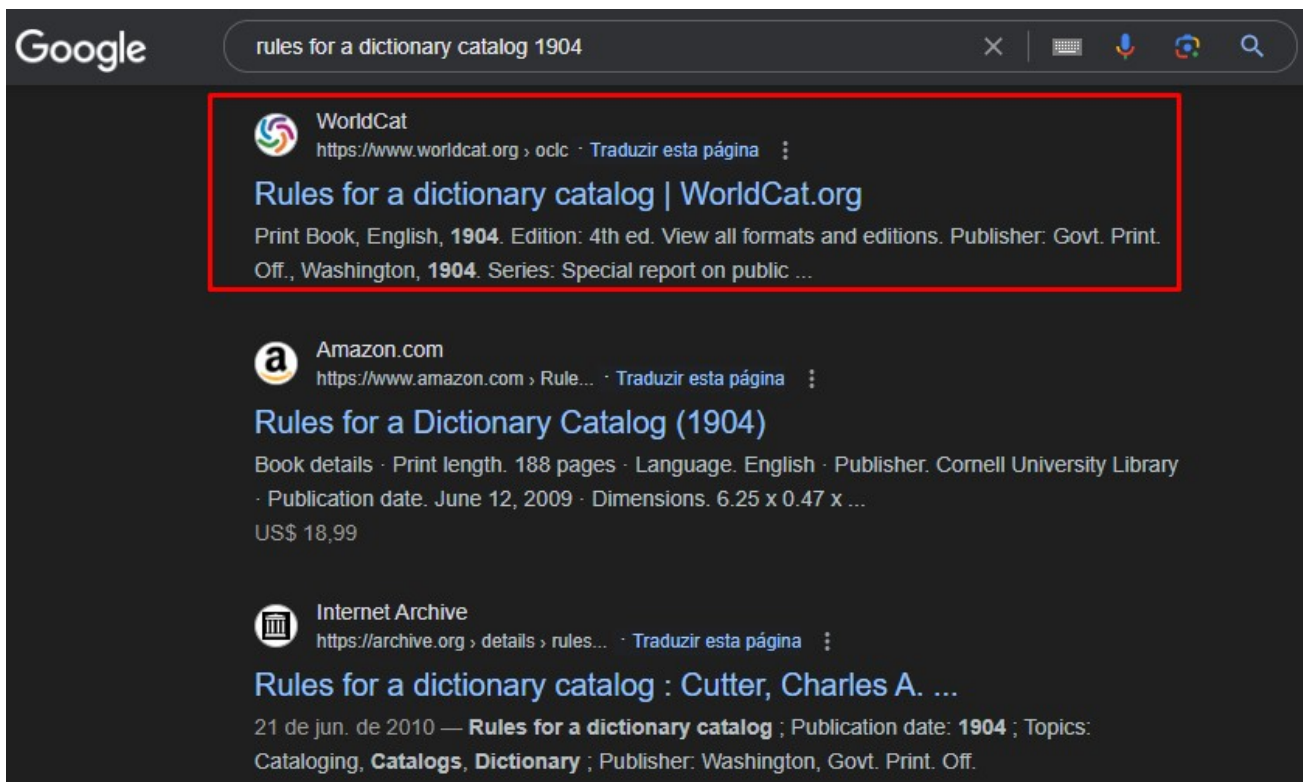


Fig. 2. WorldCat record retrieved as a search result.



Fig. 3. View source Worldcat.

and Schema.org for generic metadata. Complementing this, Gagnon et al. (2024) demonstrated a practical implementation where Schema.org functions as a coarse-grained model alongside the fine-grained IFLA LRM, specifically handling data types not natively covered by LRM, such as literary awards and user reviews. It is therefore considered that,

just as the initiatives mentioned applied Schema.org in conjunction with the FRBR model, a similar approach can be applied to the IFLA LRM. Since IFLA LRM is more suitable for the Semantic Web, its elements may be mapped to the types and properties of Schema.org.

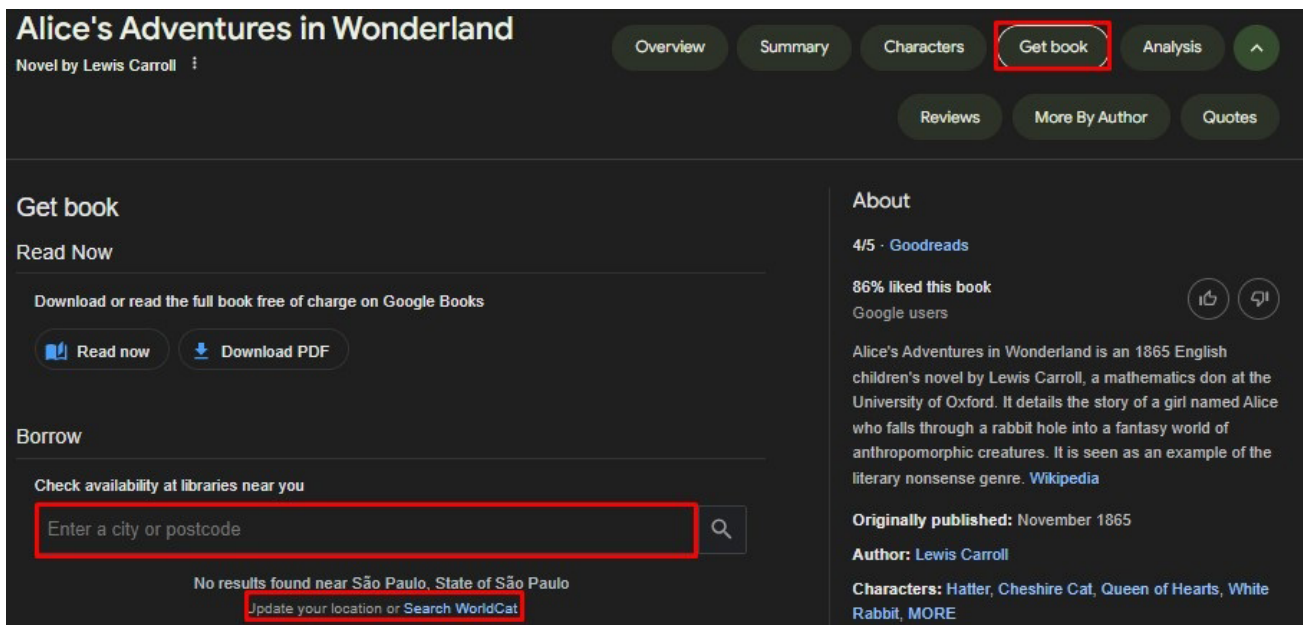


Fig. 4. Rich result for Alice’s Adventures in Wonderland book.

## 5. Crosswalk Between IFLA LRM Attributes and Schema.org Properties

In the IFLA LRM, a Work refers to the intellectual or artistic content of a distinct creation. In Schema.org, the CreativeWork class encompasses distinct works, including properties that may characterize specific instances of the main IFLA LRM entities. According to Riva et al. (2017), the relationships among the entities Work, Expression, Manifestation, and Item are essential to the model, while attributes and other relationships are not strictly required for implementation.

The IFLA LRM provides a definition for each entity and attribute, each identified by an ID. Similarly, Schema.org supplies descriptions for types and properties, while also suggesting additional properties and providing examples of codings in the recommended formats. To illustrate possible relationships between the main IFLA LRM entities and Schema.org types, a preliminary mapping was carried out, suggesting alignments between the classes and attributes/properties of the respective models.

Table 2 demonstrates a potential correlation between the primary IFLA LRM entities and their attributes with the Schema.org CreativeWork type and its associated properties. It should be noted that the figure does not display all the attributes and properties of the models but instead provides a few examples. The association among the elements was established intuitively, based on the descriptions provided by Riva et al. (2017) and Schema.org (2025d).

The alignment presented in Table 2 illustrates a possible correspondence between the attributes of the main IFLA LRM entities (Work, Expression, Manifestation, and Item) and the properties of Schema.org’s *CreativeWork*. Both direct and approximate correspondences were identified,

depending on the level of conceptual granularity of each model.

The attribute Work Category (LRM-E2-A1) was associated with the *genre* property, since within the scope of this attribute lies the category that indicates the type or genre to which the work belongs, thereby characterizing its form or style. The crosswalk reveals distinctions in the granularity of the models. For instance, the mapping of the Work Representative Expression (LRM-E2-A2) attribute requires a nuanced interpretation. In IFLA LRM, this attribute acts as a ‘container’ that records various values such as language, key, or scale derived from an expression deemed essential to characterize the work.

Therefore, mapping it to Schema.org properties like *inLanguage* or *audience* represents a partial or derived alignment. It is not a direct equivalence between the attribute and the property, but rather a mapping of a specific sub-value contained within the LRM attribute to a corresponding Schema.org property. In contrast, attributes at the Expression level, such as *Language* (LRM-E3-A6), allow for direct alignment with *inLanguage*, as both specifically describe the linguistic content of the realization.

At the level of the Expression entity, the attribute *Category* (LRM-E3-A1) was aligned with *encodingFormat*, which describes the technical type or format of the expression. The attribute *Extent* (LRM-E3-A2), which quantifies the extent of an expression, finds a parallel in *timeRequired*, a property that specifies the average time needed to consume the content. The attribute *Intended audience* (LRM-E3-A3) corresponds directly to the *audience*, while *Use rights* (LRM-E3-A4) is associated with the license describing the applicable terms of use or restrictions. The attribute *Language* (LRM-E3-A6) directly corresponds to *inLanguage*, which specifies the language of the expression.

**Table 2. Crosswalk between IFLA LRM attributes and Schema.org CreativeWork properties.**

IFLA LRM entity	IFLA LRM attribute	IFLA LRM definition	CreativeWork property	Alignment level	Definition
LRM-E2-A1 Work	Category	A type to which the work belongs.	genre	Partial	Genre of the creative work, broadcast channel or group.
LRM-E2-A2 Work	Representative expression	An attribute essential for characterizing the work, based on a representative or canonical expression.	inLanguage	Partial	The language of the content or performance or used in an action.
LRM-E2-A2 Work	Representative expression	An attribute essential for characterizing the work, based on a representative or canonical expression.	audience	Partial	An intended audience, i.e., a group for whom something was created.
LRM-E3-A1 Expression	Category	A type to which the expression belongs.	encodingFormat	Partial	Media type typically expressed using a MIME format, e.g., application/zip for a SoftwareApplication binary, audio/mpeg for .mp3 etc.
LRM-E3-A2 Expression	Extent	A quantification of the extent of the Expression.	timeRequired	Partial	Approximate or typical time it usually takes to work with or through the content of this work for the typical or target audience.
LRM-E3-A3 Expression	Intended audience	A class of users for which the expression is intended.	audience	Direct	An intended audience, i.e., a group for whom something was created.
LRM-E3-A4 Expression	Use rights	A class of use restrictions to which the expression is submitted.	license	Partial	A license document that applies to this content, typically indicated by URL.
LRM-E3-A6 Expression	Language	A language used in the expression.	inLanguage	Direct	The language of the content or performance or used in an action.
LRM-E4-A1 Manifestation	Category of carrier	A type of material to which all physical carrier.	material	Partial	A material that something is made from, e.g., leather, wool, cotton, paper.
LRM-E4-A2 Manifestation Extent	Extent	A quantification of the extent on a physical carrier.	materialExtent	Partial	The quantity of the materials being described or an expression of the physical space they occupy.
LRM-E4-A3 Manifestation	Intended audience	A class of users for which the physical Carriers are intended.	audience	Direct	An intended audience, i.e., a group for whom something was created.
LRM-E4-A5 Manifestation	Access conditions	Information as to how any of the carriers are likely to be obtained.	conditionsOfAccess	Direct	Conditions that affect the availability of, or method(s) of access to, an item.
LRM-E5-A2 Item	Use rights	A class of use and/or access restrictions to which the item is submitted.	usageInfo	Partial	This property is applicable both to works that are freely available and to those that require payment or other transactions. It can reference additional information, e.g., community expectations on preferred linking and citation conventions, as well as purchasing details. For something that can be commercially licensed, usageInfo can provide detailed, resource-specific information about licensing options.

For the Manifestation entity, the attribute Category of carrier (LRM-E4-A1) covers both the physical material of the manifestation and the medium used to record the content. For this reason, it may be mapped in Schema.org to either material (when referring to physical substances, such as paper or plastic) or to *encodingFormat* (when referring to media or digital formats, such as PDF, MP3, or analog/digital carriers). The attribute Extent (LRM-E4-A2) corresponds to *materialExtent*, but depending on the type of resource, it can be better specified through complementary properties: *numberOfPages* for textual resources, *duration* for audiovisual or musical manifestations, and *contentSize* for digital resources. These properties serve as specializations of the general notion of extent, enabling greater precision in representation. The attribute Intended audience (LRM-E4-A3) maintains its direct correspondence with audience, while Access conditions (LRM-E4-A5) were mapped to *conditionsOfAccess*, which describe the availability of, or methods of access to, the resource.

Finally, at the Item level, the attribute Location (LRM-E5-A1) could be represented in Schema.org by the property *holdingArchive*, defined in the *ArchiveComponent* class, a subclass of *CreativeWork*. This correspondence would be appropriate since it identifies the institution or collection responsible for holding the item. However, in Table 2, this property was not included, as the analysis was limited to properties documented directly under *CreativeWork*, without considering its subclasses. The attribute Use rights (LRM-E5-A2) was aligned with *usageInfo*, which provides additional details about terms of use, access, or licensing options.

Overall, the analysis highlights important differences between the models. While the LRM works with attributes that are more abstract and conceptually flexible, Schema.org offers more pragmatic properties, designed to describe resources in digital environments. This results in situations where the correspondence is direct, others where the attribute must be broken down into multiple properties, and still others where there is no exact equivalence within *CreativeWork*. Furthermore, the recurrence of alignments with the *audience* property reflects both the broad applicability of this attribute in the LRM and a certain conceptual overlap in the mapping, since the meaning of “intended audience” varies depending on whether it refers to a work, an expression, or a manifestation.

A critical challenge in this alignment is the disparity in semantic granularity regarding IFLA LRM’s Work and Expression. While LRM separates intellectual content (Work) and its realization (Expression) into distinct conceptual levels, Schema.org’s *CreativeWork* flattens this hierarchy. Structurally, *CreativeWork* subsumes attributes from both entities rather than equating to them individually. For instance, LRM defines ‘Subject’ at the Work level and ‘Language’ at the Expression level, whereas Schema.org includes both *about* and *inLanguage* as direct properties of a

single *CreativeWork*. Thus, the hierarchical distinction is replaced by property differentiation, sacrificing structural granularity but preserving functional discovery capabilities.

While this crosswalk offers a valuable basis for analysis, it should also be regarded as a preliminary and exploratory exercise. Its correspondences highlight directions for interoperability, but they remain open to refinement and validation in future applications. In this sense, the adoption of subclass properties, such as *holdingArchive*, or the use of specific properties in appropriate contexts, such as *duration* or *numberOfPages*, could enhance the semantic accuracy of the alignment. Thus, the exercise of correspondence not only demonstrates the feasibility of integration but also reveals opportunities for refinement and possible directions for developing models that are more sensitive to the specificities of the bibliographic domain.

## 6. Conclusions

As discussed in this article, with the continuous advancement of Information and Communication Technologies (ICTs), adapting catalogs to contemporary needs is crucial. The integration of semantic vocabularies, such as Schema.org, with conceptual models specific to the bibliographic domain, such as IFLA LRM, can significantly transform access to and use of catalog data.

The IFLA LRM provides a robust foundation for data modeling, enabling alignment with metadata standards and semantic vocabularies. Although Schema.org was not specifically designed for bibliographic data, its flexible structure facilitates integration and the discovery of information on the web. In this way, semantic enrichment can improve the user experience, increasing the likelihood of interaction with catalog resources and, consequently, promoting more effective use of the data. In conclusion, the proposed alignment between IFLA LRM and Schema.org illustrates a promising direction for semantic interoperability in library catalogs. Future studies should explore practical implementations, test crosswalks in real cataloging systems, and assess their impact on retrieval, discoverability, and user interaction.

It should be emphasized that the relationships between IFLA LRM entities and attributes and the types and properties of Schema.org are not fully compatible, as the models clearly pursue different objectives. Nevertheless, both IFLA LRM and Schema.org are flexible, particularly in relation to attributes, allowing for adaptations when necessary. Ultimately, fostering interoperability between IFLA LRM and Schema.org is not only a technical challenge but also an opportunity for libraries to assert their role in the global knowledge ecosystem of the Semantic Web.

The results observed in OCLC’s work highlight the effectiveness of applying semantic technologies in the library context, providing examples that may inspire and guide the modernization of services and information processing. Thus, the investigations conducted so far provide an initial foundation for integrating the IFLA LRM and Schema.org

models, enabling more in-depth analyses of their association, feasibility, and implications. The ongoing dialogue between bibliographic conceptual models and web vocabularies will be key to ensuring that libraries remain central actors in the evolving knowledge infrastructure of the Semantic Web.

## Availability of Data and Materials

All data reported in this study are available within the article.

## Author Contributions

ACSA: designed the research study and provided help and advice. DOFM: performed the research and wrote the manuscript. Both authors contributed to editorial changes in the manuscript. Both authors read and approved the final manuscript. Both authors participated sufficiently in the work and agreed to be accountable for all aspects of it.

## Acknowledgment

We gratefully acknowledge the support of the National Council for Scientific and Technological Development (CNPq) and the Graduate Program in Information Science at the Federal University of São Carlos (PPGCI/UFSCar).

## Funding

This research was funded by the National Council for Scientific and Technological Development (CNPq) (grant number: 308412/2022-2).

## Conflicts of Interest

The authors declare no conflicts of interest.

## Declaration of AI and AI-Assisted Technologies in the Writing Process

During the preparation of this work, the authors used artificial intelligence (AI) tools to check spelling and grammar. After using these tools, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

## References

- Amazon. CreativeWork Intents. 2023. Available at: <https://developer.amazon.com/ja-JP/docs/alexa/custom-skills/creativework-intents.html> (Accessed: 7 December 2024).
- Arakaki AC. O modelo IFLA Library Reference Model e o linked data. *Informação & Informação*. 2020; 25: 163–186. <https://doi.org/10.5433/1981-8920.2020v25n3p163>. (In Portuguese)
- Ballinger L. Metadata through the pages of Information Standards Quarterly. National Information Standards Organization. 2015. Available at: <https://www.niso.org/niso-io/2015/09/metadata-through-pages-information-standards-quarterly> (Accessed: 7 December 2024).

Bibliograph.net. BiblioGraph.net Vocabulary: deprecated. 2019. Available at: <https://bibliograph.github.io/BiblioGraph-Frozen/bibliograph.net/vocabularydeprecated.html>. (Accessed: 7 December 2024).

European Union. Linking data: What does it mean? 2022. Available at: <https://data.europa.eu/en/publications/datastories/linking-data-what-does-it-mean> (Accessed: 7 December 2024).

Fons T, Penka J, Wallis R. OCLC's Linked Data initiative: Using Schema.org to make library relevant on the Web. *Information Standards Quarterly*. 2012; 24: 29–33. <http://doi.org/10.3789/isqv24n2-3.2012.05>.

Gagnon M, Font L, Zouaq A. An Exploration of IFLA LRM for Literature Data Representation. *ACM Journal on Computing and Cultural Heritage*. 2024; 17: 1–21. <https://doi.org/10.1145/3687486>.

Godby CJ. The Relationship between BIBFRAME and OCLC's Linked-Data Model of Bibliographic Description: A Working Paper. OCLC Research: Dublin, Ohio. 2013.

Godby CJ, Denenberg R. Common ground: Exploring compatibilities between the Linked Data models of the Library of Congress and OCLC. 2015a. Available at: <http://www.oclc.org/content/dam/research/publications/2015/oclcresearch-linked-data-2015.pdf> (Accessed: 7 December 2024) (preprint)

Godby CJ. A division of labor: the role of Schema.org in a Semantic Web model of library resources. 2016. Available at: <https://www.oclc.org/research/publications/2016/schema-role-semantic-web-library.html>. (Accessed: 7 December 2024).

Godby CJ, Wang S, Mixter JK. Library linked data in the cloud: OCLC's experiments with new models of resource description. *Synthesis Lectures on the Semantic Web: Theory and Technology*. 2015b. Available at: <https://www.oclc.org/research/publications/2015/oclcresearch-library-linked-data-in-the-cloud.html> (Accessed: 7 December 2024).

Google. How Google's Knowledge Graph works. 2024. Available at: <https://support.google.com/knowledgepanel/answer/9787176?hl=en> (Accessed: 7 December 2024).

Google Search Central. Book actions (Book) structured data. 2025. Available at: <https://developers.google.com/search/docs/appearance/structured-data/book?hl=pt-br#create-your-feed> (Accessed: 7 December 2024).

Guha R, Brickley D, Macbeth S. Schema.org: Evolution of structured data on the Web. *Communications of the ACM*. 2016; 59: 44–51. <https://doi.org/10.1145/2844544>.

International Federation of Library Associations and Institutions (IFLA). Statement of International Cataloguing Principles (ICP). IFLA: Den Haag. 2016.

Iliadis A, Acker A, Stevens W, Kavakli SB. One schema to rule them all: How Schema.org models the world of search. *Journal of the Association for Information Science and Technology*. 2025; 76: 460–523. <https://doi.org/10.1002/asi.24744>.

LIBER Linked Open Data Working Group. Best Practices for Li-

- brary Linked Open Data (LOD) Publication. 2021. Available at: <https://libereurope.eu/wp-content/uploads/2021/02/LOD-Guidelines-FINAL-Feb-2021.pdf> (Accessed: 7 December 2024). (pre-print)
- Meta. Tags de microdados. 2025. Available at: <https://developers.facebook.com/docs/marketing-api/catalog/guides/microdata-tags> (Accessed: 7 December 2024).
- National Library Board. About us. 2025. Available at: <https://www.nlb.gov.sg/main/about-us> (Accessed: 7 December 2024).
- OCLC. Rules for a dictionary catalog. 2024a. Available at: <https://search.worldcat.org/pt/title/367298848> (Accessed: 7 December 2024).
- OCLC. View-source. 2024b. Available at: <https://search.worldcat.org/pt/title/367298848> (Accessed: 7 December 2024).
- Pinterest. Create rich Pins. 2023. Available at: <https://help.pinterest.com/en/business/article/rich-pins> (Accessed: 7 December 2024).
- Riva P, Le Boeuf P, Žumer M. IFLA Library Reference Model: A conceptual model for bibliographic information. IFLA: Den Haag. 2017.
- Schema.org. Schema.org. 2025a. Available at: <https://schema.org> (Accessed: 7 December 2024).
- Schema.org. Schema.org extensions. 2025b. Available at: <https://schema.org/docs/extension.html> (Accessed: 7 December 2024).
- Schema.org. Data model. 2025c. Available at: <https://schema.org/docs/datamodel.html> (Accessed: 7 December 2024).
- Schema.org. CreativeWork. 2025d. Available at: <https://schema.org/CreativeWork> (Accessed: 7 December 2024).
- Wallis R. From Ambition to go live. 2023. Available at: [https://swib.org/swib23/slides/02\\_Richard%20Wallis\\_\\_From%20Ambition%20to%20Go%20Live%20SWIB.pdf](https://swib.org/swib23/slides/02_Richard%20Wallis__From%20Ambition%20to%20Go%20Live%20SWIB.pdf). (Accessed: 7 December 2024).
- World Wide Web Consortium. Schema Bib Extend Community Group. 2025. Available at: <https://www.w3.org/community/schemabibex> (Accessed: 7 December 2024).
- Zafalon ZR. Recurso informacional e representação documental. Encontro de Representação Documental – EnReDo. CPOI/UFSCar: São Carlos. 2017. (In Portuguese)