

CONECTANDO DADOS BIBLIOGRÁFICOS À WEB: IFLA LIBRARY REFERENCE MODEL E SCHEMA.ORG

Connecting Bibliographic Data to the Web: IFLA Library Reference Model and Schema.org

Dayane Onaga Ferreira Machado

Mestre em Ciência da Informação. Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT), Brasília, Brasil.
onagadayane@gmail.com. <https://orcid.org/0000-0001-5772-418X>.

Ana Carolina Simionato Arakaki

Pós Doutorado em Ciência da Computação – Universidade de São Paulo (USP), São Paulo, Brasil. Doutorado em Ciência da Informação – Universidade Estadual Paulista (UNESP), Marília, São Paulo, Brasil. anaarakaki@ibict.br.
<https://orcid.org/0000-0002-0140-9110>.

RESUMO

Objetivo: As bibliotecas disponibilizam informações relevantes e confiáveis por meio dos catálogos digitais. Contudo, nos catálogos tradicionais, os dados não são conectados e compartilhados na *Web*. Desse modo, é necessário revitalizá-los, utilizando modelos conceituais, como o IFLA *Library Reference Model* (LRM), e formatos semânticos para interoperabilidade na *Web*. Assim, esta pesquisa objetiva discutir a modelagem dos catálogos orientada pelo LRME a utilização do *Schema.org*. **Método:** Esta pesquisa é de abordagem qualitativa e do ponto de vista dos objetivos, é exploratória e descritiva. **Conclusões:** O LRM orienta a modelagem considerando o alinhamento com padrões no âmbito do *Linked Open Data*. Nesse contexto, o *Schema.org* é um vocabulário semântico viável, que pode ser aplicado ao domínio bibliográfico, permitindo a integração dos dados dos catálogos à *Web*.

Palavras-chave: Catálogos digitais. IFLA Library Reference Model. Schema.org. Enriquecimento Semântico.

ABSTRACT

Objective: Libraries provide relevant and reliable information through digital catalogs. However, in traditional catalogs, data are not connected or shared on the *Web*. Thus, it is necessary to revitalize them by using conceptual models, such as the IFLA *Library Reference Model* (LRM), and semantic formats for interoperability on the *Web*. This research, therefore, aims to discuss catalog modeling guided by the LRM and the use of *Schema.org*. **Method:** This study adopts a qualitative approach and, in terms of objectives, is exploratory and descriptive. **Conclusions:** The LRM guides modeling by considering alignment with standards within the scope of *Linked Open Data*. In this context, *Schema.org* is a viable semantic vocabulary that can be applied to the bibliographic domain, allowing for the integration of catalog data with the *Web*.

keywords: Digital Catalog. IFLA Library Reference Model. Schema.org. Semantic enrichment.

1 INTRODUÇÃO

No âmbito das bibliotecas, os catálogos digitais são importantes instrumentos de busca e recuperação de recursos informacionais, se destacando como fontes confiáveis e relevantes. No entanto, os métodos utilizados para tratar essas informações estão desatualizados frente às novas tecnologias e à forma como a informação é atualmente disponibilizada, especialmente na *Web*.

De modo geral, nos catálogos não há o estabelecimento de relacionamentos entre os registros bibliográficos e tão pouco com fontes externas. Nesse sentido, Serra e Santarém Segundo (2021) destacam que, ao replicar ou criar registros para seus catálogos, as instituições acabam formando um silo de dados. Os autores observam que a ausência de conexões entre os próprios itens da coleção intensifica esse problema, gerando “silos dentro do silo” (Serra; Santarém Segundo, 2021). Certamente, esse cenário é desfavorável para as bibliotecas, visto que há um empenho significativo no tratamento dos recursos informacionais para criação de metadados de qualidade. Em vista disso, para atender as necessidades e expectativas dos usuários contemporâneos e favorecer o uso efetivo dos catálogos, é fundamental que esses instrumentos sejam atualizados e adaptados às novas demandas.

Nesse contexto, os modelos conceituais desempenham um papel fundamental na estruturação eficiente de sistemas, fornecendo uma base sólida para a organização de estruturas complexas. O IFLA *Library Reference Model* (LRM) é um modelo conceitual de alto nível, que objetiva orientar a formulação de regras de catalogação e a implementação de sistemas bibliográficos (Riva; Bœuf; Žumer, 2017). O modelo fornece uma estrutura conceitual abrangente e propõe facilitar a reutilização dos dados, levando em consideração aplicações em *Linked Open Data* (LOD). Arakaki (2020) aponta que a instanciação do modelo e a possibilidade de criar *namespaces* podem favorecer o mapeamento com padrões de metadados.

Nessa perspectiva, a utilização de vocabulários semânticos é essencial para permitir a integração e o aproveitamento dos dados na *Web*. Para atingir esse objetivo, uma alternativa é o uso do *Schema.org*, um vocabulário semântico desenvolvido pelos principais mecanismos de busca. Os dados estruturados com

esse vocabulário são enriquecidos, facilitando a interpretação, a recuperação dos registros e a conexão com outras fontes de informações.

Embora vários estudos tenham abordado diferentes aspectos da modernização dos catálogos, ainda há uma lacuna significativa na literatura quanto à associação entre o LRM e vocabulários semânticos, como o *Schema.org*. Em vista disso, a partir de uma abordagem qualitativa e de uma pesquisa descritiva e exploratória, este artigo objetiva discutir sobre a modelagem dos dados bibliográficos pelo viés do modelo LRM, assim como sobre a utilização do *Schema.org* para revitalização dos catálogos digitais. Segundo Gil (2008) a pesquisa exploratória permite o estudo dos temas sob diversos ângulos e aspectos, este tipo de pesquisa é realizada especialmente quando os temas escolhidos são pouco explorados.

O artigo está organizado em cinco seções principais. Além desta seção de introdução que apresenta o contexto, os objetivos e a natureza da pesquisa, a segunda seção aborda o IFLA *Library Reference Model*, descrevendo as principais características do modelo. Na terceira seção, discute-se sobre os principais aspectos do *Schema.org*. A quarta seção trata dos resultados aprimorados, discorrendo sobre como os dados estruturados podem gerar resultados enriquecidos. Por fim, a quinta seção traz as considerações finais.

2 IFLA LIBRARY REFERENCE MODEL

O *IFLA Library Reference Model* (LRM) é um modelo de alto nível do tipo entidade-relacionamento, desenvolvido a partir dos *FunctionalRequirements* (FR). Esses incluem os modelos conceituais: *FunctionalRequirements for BibliographicRecords* (FRBR) focado em registros bibliográficos, *FunctionalRequirements for Authority Data* (FRAD) especializado em registros de autoridade, e o *FunctionalRequirements for SubjectAuthority Data* (FRSAD) para registros de assunto (Riva; Bœuf; Žumer 2017). Os modelos partem de pontos de vista distintos, assim, no LRM os aspectos relevantes foram harmonizados, resultando em uma estrutura conceitual mais coesa e abrangente.

Conforme explanam Riva, Bœuf e Žumer (2017) o escopo funcional do LRM é baseado em tarefas genéricas executadas pelos usuários finais e intermediários,

sendo: encontrar, identificar, selecionar, obter e explorar. A tarefa ‘encontrar’ refere-se ao atendimento dos critérios de busca; ‘identificar’ diz respeito à possibilidade de distinguir entre recursos semelhantes; ‘selecionar’ refere-se à adequação dos recursos encontrados em relação às preferências do usuário; ‘obter’ trata do acesso ao recurso, e, por fim, ‘explorar’ diz respeito à descoberta de recursos, a partir do contexto e dos relacionamentos entre eles.

Os principais elementos que compõem o LRM são: entidades, atributos e relacionamentos. As entidades caracterizam-se como classes, os atributos conferem sentido às instâncias das entidades e os relacionamentos representam a conexão entre elas, contextualizando-as (Riva; Bœuf; Žumer, 2017). O Quadro 1 apresenta as entidades do LRM, as hierarquias e as definições.

Quadro 1 - Entidades, hierarquia e definições LRM

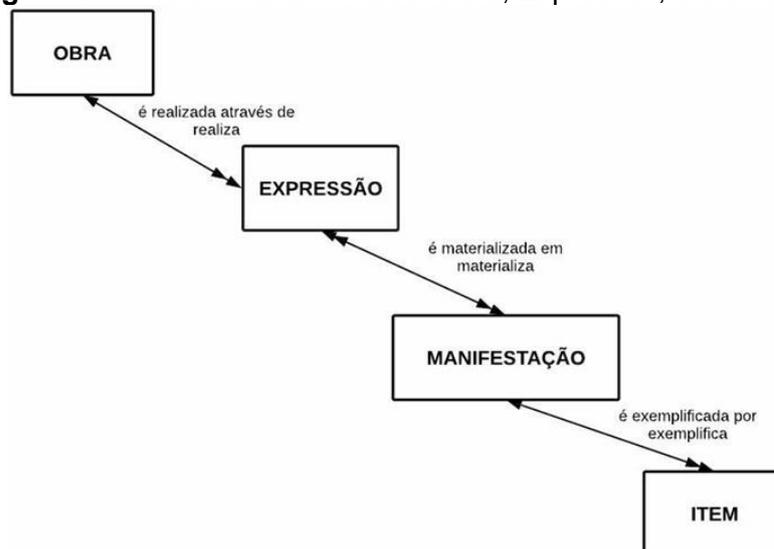
Entidade	Hierarquia	Definição
LRM-E1: Res	Nível superior	Qualquer entidade no universo do discurso.
LRM-E2: Obra	Segundo nível	Uma distinta criação intelectual ou artística.
LRM-E3: Expressão	Segundo nível	Uma combinação específica de sinais que transmitem o conteúdo intelectual ou artístico.
LRM-E4: Manifestação	Segundo nível	Um conjunto de todos os suportes que supostamente compartilham as mesmas características do conteúdo intelectual ou artístico e aspectos da forma física.
LRM-E5: Item	Segundo nível	Um ou vários objetos que carregam sinais destinados a transmitir conteúdo intelectual ou artístico.
LRM-E6: Agente	Segundo nível	Uma entidade capaz de ações deliberadas, de cessão de direitos e de ser responsabilizada por suas ações.
LRM-E7: Pessoa	Terceiro nível	Um ser humano individual.

LRM-E8: Agente coletivo	Terceiro nível	Uma reunião ou organização de pessoas com um nome específico e capaz de atuar como uma unidade.
LRM-E9: Nomen	Segundo nível	Uma associação entre uma entidade e uma designação.
LRM-E10: Lugar	Segundo nível	Uma extensão de espaço determinada.
LRM-E11: Intervalo de tempo	Segundo nível	Uma extensão temporal que possui um início, um fim e uma duração

Fonte: Riva, Bœuf e Žumer (2017).

No Quadro 1 são apresentadas todas as entidades do modelo. A entidade superior é Res, que têm como subclasses: Obra, Expressão, Manifestação, Item, Agente, Pessoa, Agente Coletivo, Nomen, Lugar e Intervalo de Tempo. As relações entre as principais entidades do modelo são ilustradas na Figura 1.

Figura 1 - Relacionamentos entre: Obra, Expressão, Manifestação e Item



Fonte: Riva; Bœuf e Žumer (2017, p. 85).

Como ilustra a figura acima, a cardinalidade é indicada pela ponta das setas, sendo que uma ponta significa que a cardinalidade é um (1) e uma seta de pontas duplas indica que a cardinalidade é muitos (M). Desse modo, compreende-se que,

uma obra é realizada por meio de uma ou mais expressões; uma expressão é a realização de apenas uma obra; uma expressão pode ser incorporada em mais de uma manifestação, que pode ser exemplificada por um ou mais itens, representados apenas por uma manifestação. Em outras palavras, uma obra é a criação intelectual ou artística idealizada por um indivíduo, ou organização, a expressão é a realização desta obra, podendo ser expressa em diferentes formatos e idiomas; a manifestação é a materialização da expressão em um ou mais itens em formato físico, ou digital, por exemplo.

O LRM foi projetado para o contexto da *Web Semântica*, visando fornecer significado aos dados e facilitar a interpretação pelos sistemas computacionais. Dessa maneira, o modelo permite implementar sistemas bibliográficos mais eficientes, reduzindo a ambiguidade, e possibilitando a atribuição de *namespaces* para aplicações de *Linked Open Data* (LOD) (Riva; Bœuf; Žumer, 2017).

Além da vinculação dos dados, é essencial considerar a descoberta desses dados na *Web*. A Declaração dos Princípios Internacionais de Catalogação (PIC), destaca, no princípio 2.10, a utilização do uso de vocabulários que facilitem a tradução automática e a desambiguação para promover o intercâmbio de dados e ferramentas de descoberta (Galeffi *et al.*, 2016).

Desse modo, a conexão dos dados e uso de vocabulários semânticos, promovem contexto ao processo de descrição e criam conexões que podem levar à descoberta acidental de informações relevantes, ao mesmo tempo, em que agilizam a recuperação. Para isso, a padronização dos dados conforme vocabulários reconhecidos pelos mecanismos de busca é essencial. Nessa perspectiva, a próxima seção discorre sobre o *Schema.org*, um vocabulário semântico que promove a descoberta dos dados na *Web*.

3 SCHEMA.ORG

O *Schema.org* foi desenvolvido pelos mecanismos de busca *Google*, *Microsoft*, *Yahoo* e *Yandex*. Trata-se de uma iniciativa colaborativa e comunitária, com o objetivo de criar, manter e disponibilizar um conjunto de esquemas para dados estruturados, visando otimizar a interpretação e a recuperação de

informações pelos mecanismos de busca na *Web* (Schema.org, 2024a). Atualmente o vocabulário está na versão 28.0 e tem sido adotado amplamente na *Web*. Conforme indica o site oficial do *Schema.org*, em 2024, estima-se que mais de 45 milhões de páginas aderiram ao conjunto de esquemas. Segundo *Schema.org* (2024a) a adoção de um vocabulário comum facilita a escolha de esquemas e a obtenção dos melhores resultados para *webmasters* e desenvolvedores.

O vocabulário é composto por um conjunto de 'tipos' ou 'classes', associadas a um conjunto de propriedades. A Figura 2 ilustra uma parte da hierarquia principal, apresentando os desdobramentos das classes *Thing* e *CreativeWork*.

Figura 2 - Hierarquia principal do *Schema.org*



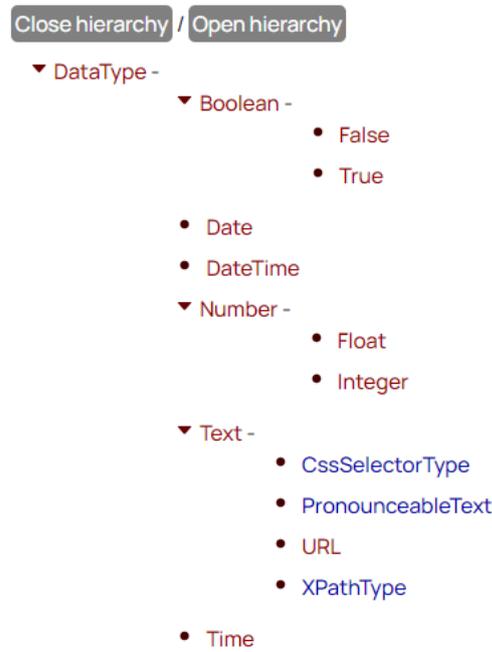
Fonte: *Schema.org* (2024c).

Esses elementos são organizados hierarquicamente, com *Thing* sendo a classe de nível superior, que inclui as subclasses: *Action*, *BioChemEntity*, *CreativeWork*, *Event*, *Intangible*, *MedicalEntity*, *Organization*, *Person*, *Place*, *Product* e *Taxon*. Todas as classes herdam propriedades da superclasse e possuem suas próprias subclasses e propriedades específicas. Há também uma hierarquia paralela

com os tipos de dados, que define os tipos de valores que podem ser especificados para as propriedades. A Figura 3 ilustra a hierarquia para os tipos de dados.

Figura 3 - Hierarquia de tipos de dados

DataTypes:



Fonte: Schema.org (2024c).

A hierarquia inclui os seguintes tipos de dados: *boolean*, com os valores *false* e *true*, que expressam dois tipos de condições ou estados; *date*, que inclui datas no formato ISO 8601 (ano-mês-dia); *datetime*, que representa a combinação de data e hora; *number*, para valores numéricos, que incluem tipos específicos como *float* e *integer*; *text*, que representa valores em formato de texto e *time*, utilizado para expressar horários específicos.

Hitzler (2021) destaca que o *Schema.org*, assim como, aplicações de modelagem de ontologias e o *Wikidata*, surgiram no âmbito da *Web Semântica*, com o objetivo de facilitar o compartilhamento, a descoberta, a integração e a reutilização de dados. Sendo assim, o *Schema.org* pode ser considerado um dos exemplos de resultado proveniente dos diversos estudos sobre a concretização da *Web Semântica*. O vocabulário é compatível com formatos semânticos, recomendados pelo *World Wide Web Consortium (W3C)*, como: *ResourceDescription Framework in Attributes (RDFa)*, *Microdados* e *JavaScriptObjectNotation for Linking Data (JSON-*

LD). A utilização desses formatos possibilita que os mecanismos de busca compreendam melhor o conteúdo e o contexto das informações, proporcionando experiências de busca mais satisfatórias.

4 RESULTADOS APRIMORADOS

Os dados estruturados são fundamentais para que os mecanismos de busca possam coletar, interpretar e classificar o conteúdo disponível na *Web*, o *Google*, por exemplo, recomenda o uso do *Schema.org*. Muitos dos elementos do vocabulário não são essenciais para pesquisa no *Google*, mas podem ser úteis para outros mecanismos, serviços e plataformas (Google Search Central, 2024b).

O *Schema.org* facilita a integração e a abertura de dados na *Web*, permitindo, por exemplo, a inclusão de dados estruturados no *KnowledgeGraph*. Este banco de dados organiza e relaciona fatos sobre entidades, utilizando informações de fontes confiáveis (Google, 2024). Desse modo, os dados podem ser recuperados como fragmentos enriquecidos, denominados “*richsnippets*”, isto é, podem ser recuperados como resultados enriquecidos com informações adicionais e elementos visuais. Essa abordagem não só melhora a qualidade das informações, mas também proporciona uma experiência mais rica e informativa, oferecendo ao usuário informações contextuais relevantes, como dados históricos, biografias e conexões entre temas.

Nesse contexto, Wallis (2022) destaca que o *Schema.org* é um vocabulário de *Linked Data* aberto para a *Web*, baseado em *ResourceDescription Framework* (RDF), o que facilita a interoperabilidade e a integração dos dados estruturados em diferentes plataformas. Para que os dados sejam recuperados pelos mecanismos de busca, podem ser incorporados diretamente no código HTML de uma página, nos formatos, JSON-LD, RDFa ou Microdados. O formato JSON-LD é o mais recomendado, pois separa os dados estruturados do HTML, facilitando a manutenção e o desenvolvimento. Já o RDFa e os Microdados integram as informações diretamente no código HTML por meio de atributos específicos (Google Search Central, 2024b).

Por ser um vocabulário genérico, o *Schema.org* pode ser utilizado em diversos contextos, como evidenciado por sua aplicação em plataformas distintas: o *Pinterest* aplica *tags* para implementar ‘*Rich Pins*’ (Pinterest, 2024); o *Facebook* enriquece *links* compartilhados (Meta, 2024); a *Amazon* emprega dados estruturados para interpretar e responder às consultas dos usuários por meio da *Alexa* (Amazon, 2023). No âmbito acadêmico, pesquisas recentes têm explorado o vocabulário para representação de dados de pesquisa (Wu *et al.*, 2022; Payne; Verhey, 2022; Alrashed *et al.*, 2021; Silva, 2020), na área de *e-commerce* (Belz, 2022; Kejriwal *et al.*, 2021; Peeters *et al.*, 2020) e para redes sociais (Jones *et al.*, 2021; Tomoyse; Santos; Simionato, 2019; García; Gil, 2019), entre outros campos de estudo.

No domínio bibliográfico, o *Schema.org* tem sido aplicado para o enriquecimento semântico dos dados de catálogos, conforme evidenciam as pesquisas de Wallis (2023) e Godby (2016). A adaptação do modelo de dados utilizado nos catálogos pode estabelecer a comunicação com a *Web*. A abertura dos dados pode favorecer tanto a integração a outros conjuntos de dados nos *KnowledgeGraphs* quanto a recuperação dos registros bibliográficos diretamente na página de resultados de busca. Nessa perspectiva, a próxima seção discorre sobre o *Schema.org* no contexto das bibliotecas.

4.1 Resultados aprimorados para recursos bibliográficos

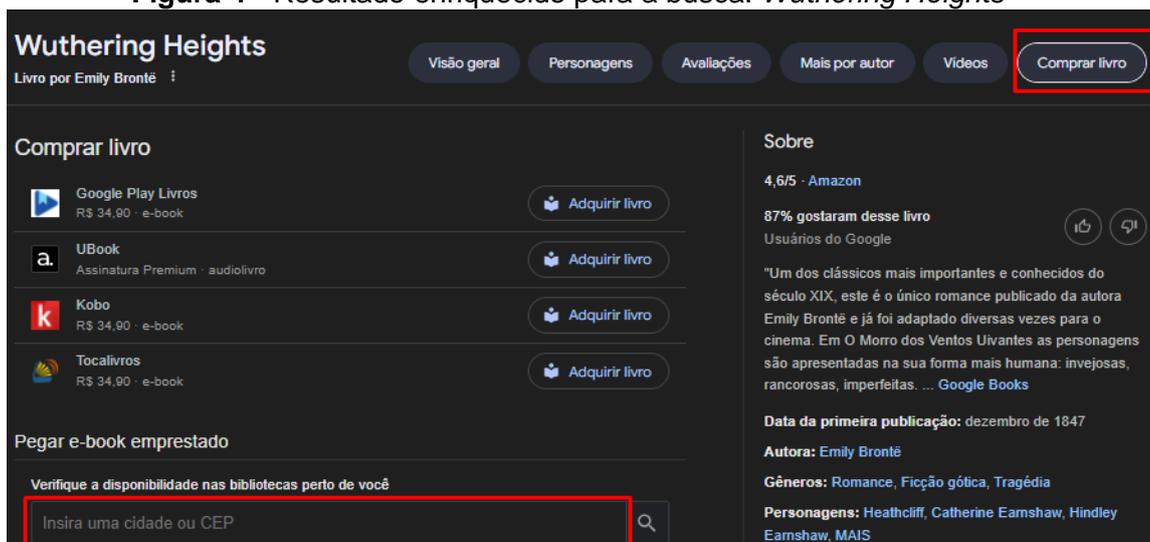
Para garantir a interoperabilidade e a integração eficiente dos dados dos catálogos na *Web*, é essencial adotar um modelo de referência apropriado. O LRM é um modelo internacionalmente reconhecido, desenvolvido tendo em vista a reutilização dos dados no contexto do *Linked Open Data* e a integração com aplicações semânticas. Riva, Bœuf e Žumer (2017) explanam que, o cenário atual se transformou desde a criação original do modelo FRBR, trazendo novas demandas, especialmente no que diz respeito à reutilização de dados em aplicações da *Web Semântica* (Riva; Bœuf; Žumer, 2017).

O modelo oferece uma abordagem detalhada para a representação de diversos tipos de recursos, estabelecendo conexões entre entidades e atributos em um domínio. Conforme ressalta Arakaki (2020), a viabilidade semântica dos atributos

depende da definição dos metadados que instanciam o modelo, ou seja, os metadados darão forma à estrutura proposta pelo IFLA LRM, que deve servir como base para a criação de instrumentos de representação, como padrões e códigos de catalogação.

Nessa perspectiva, o modelo de dados do *Schema.org* é amplo e flexível, atribuindo significado aos dados. A versatilidade do vocabulário possibilita sua aplicação em contextos distintos. Dessa maneira, se aplicado aos dados dos catálogos, os registros podem ser recuperados como resposta a uma busca no *Google*, por exemplo. Ainda é possível que os resultados apresentem as opções de compra ou empréstimo. Os dados estruturados e alinhados ao vocabulário do *Schema.org* permitem recuperar resultados semelhantes aos apresentados na Figura 4.

Figura 4 - Resultado enriquecido para a busca: *Wuthering Heights*



Fonte: Google (2024).

Como pode ser visualizado na Figura 4, foi realizada uma busca pelo livro *Wuthering Heights* da escritora Emily Brontë, a estrutura apresentada na imagem foi exibida diretamente na página de resultados. As opções disponíveis incluem: visão geral, personagens, avaliações, mais informações sobre o autor, vídeos e a opção de comprar o livro. Ainda, são exibidos a descrição do conteúdo, as avaliações, as indicações de lojas para compra e a opção de empréstimo em uma biblioteca. A opção 'empréstimo' é ativada ao selecionar 'comprar livro', disponibilizando um

campo (destacado em vermelho na figura) para que o usuário informe seu CEP e verifique se o item está disponível em uma biblioteca próxima.

Essas funcionalidades podem ser habilitadas seguindo as orientações do Google. No caso da opção de empréstimo, é necessário criar uma estrutura de dados denominada 'feed', nos formatos JSON-LD, Microdados, ou RDFa, e utilizar as propriedades definidas pelo *Schema.org* (Google Search Central, 2024a). A Figura 5 ilustra um exemplo de *feed* em JSON-LD para '*BorrowAction*', que habilita a opção de empréstimo.

Figura 5 -*Feed BorrowAction*

```
{
  "@context": "https://schema.org",
  "@type": "DataFeed",
  "dataFeedElement": [
    {
      "@context": "https://schema.org",
      "@type": "Book",
      "@id": "https://example.com/work/the_catcher_in_the_rye",
      "url": "https://example.com/work/the_catcher_in_the_rye",
      "name": "The Catcher in the Rye",
      "author": {
        "@type": "Person",
        "name": "J.D. Salinger"
      },
      "sameAs": "https://en.wikipedia.org/wiki/The_Catcher_in_the_Rye",
      "workExample": [
        {
          "@type": "Book",
          "@id": "https://example.com/edition/the_catcher_in_the_rye_paperback",
          "isbn": "9787543321724",
          "bookEdition": "Mass Market Paperback",
          "bookFormat": "https://schema.org/Paperback",
          "inLanguage": "en",

```

Fonte:Google Search Central (2024a)

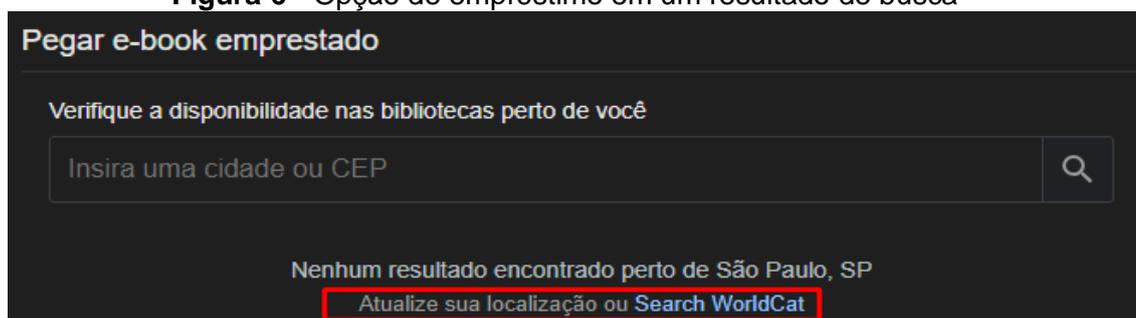
As ações relacionadas a livros possibilitam que os usuários descubram autores e livros e onde podem ter acesso, seja por meio da compra ou empréstimo. A especificação '*ReadAction*', proveniente do *Schema.org*, habilita a opção de compra, enquanto '*BorrowAction*' habilita o empréstimo (Google Search Central, 2024a). Além dessas especificações, é necessário criar um *feed* com as propriedades indicadas para entidade biblioteca e outro para entidade livro.

No contexto dos mecanismos de busca e do *Schema.org*, uma “obra” é considerada um objeto tangível que pode ser acessado e adquirido pelo usuário. Como descrito em *Schema.org* (2024b, tradução nossa) “[...] algo pode ser simultaneamente um Livro e um Produto e ser descrito de forma útil com propriedades de ambos os tipos.”

No âmbito das bibliotecas, o *Schema.org* foi implementado no catálogo *Worldcat*, gerenciado pela *Online Computer Library Center* (OCLC). Nesse contexto, o vocabulário foi alinhado às conceitualizações dos FRBR. Conforme descrito por Godby (2016), os especialistas da OCLC consideraram estruturar a classe *CreativeWork* em triplas RDF para construir um modelo semântico para os dados do *Worldcat*. A classe inclui metadados para descrição de livros, artigos, trabalhos acadêmicos, pinturas, esculturas, manuais, fotografias, gravações musicais, filmes, séries, jogos, entre outros.

É possível visualizar os resultados dos trabalhos da OCLC na prática. Os registros do catálogo podem aparecer diretamente na página de resultados de busca, enriquecidos com informações adicionais. Como os resultados são dinâmicos, a maneira como são exibidos pode variar dependendo do momento da pesquisa. A Figura 6 apresenta um recorte do mesmo formato de exibição apresentado anteriormente, na Figura 4, e inclui a recomendação (destacada em vermelho) para o usuário acessar o *WorldCat*.

Figura 6 - Opção de empréstimo em um resultado de busca



Fonte: Google (2024).

Conforme apresenta a Figura 6, ao inserir "São Paulo, SP" no campo de pesquisa, é exibida uma mensagem informando que nenhum resultado foi encontrado para o recurso buscado. No entanto, a mensagem sugere atualizar o CEP ou realizar a pesquisa pelo livro diretamente no *WorldCat*. Assim, mesmo que a

busca não retorne resultados provenientes do catálogo, o usuário é direcionado para ele, permitindo-lhe explorar o conteúdo dos acervos e descobrir outros recursos de interesse.

Ainda sobre a aplicação do Schema.org no contexto das bibliotecas, destaca-se também uma iniciativa mais recente apresentada na LD4 *Conference on Linked Data* de 2023. Nesta conferência, Richard Wallis, um dos principais contribuintes do Schema.org, expôs o projeto de estruturação de dados desenvolvido para a *National Library Board Singapore*. Durante sua apresentação, Wallis abordou como foram implementados os dados vinculados usando o *Bibframe* na *National Library Board* (NLB) - uma instituição governamental que tem como missão fomentar a leitura, desenvolver competências informacionais e estimular o conhecimento sobre a história e o patrimônio cultural de Singapura (National Library Board, 2024).

De acordo com Wallis (2023), o Schema.org desempenhou um papel fundamental no desenvolvimento do modelo de dados no projeto da NLB. A estratégia adotada envolveu o alinhamento dos dados em MARC-XML, DC-XML e CSV com o Bibframe, permitindo preservar tanto a integridade quanto o detalhamento dos registros bibliográficos. Paralelamente, o Schema.org foi implementado para a descrição das entidades, o que otimizou sua descoberta e interpretação pelos mecanismos de busca. O autor relata que os resultados obtidos evidenciaram o êxito na transformação e integração dos dados da NLB em um formato estruturado, tornando-os apropriados para análise e exploração em gráficos do conhecimento na *Web* (Wallis, 2023).

Quando um domínio é modelado com base em modelos de referência adequados, como o LRM ou FRBR, e estruturado utilizando formatos e vocabulários semânticos, como o Schema.org, possibilita-se a integração dos dados das bibliotecas aos resultados de busca aprimorados. Ao acessar essas estruturas, os mecanismos de busca podem interpretar o contexto e o significado dos dados, tornando a recuperação de informações mais precisa e relevante. Dessa forma, as bibliotecas podem conectar seus acervos a uma rede mais ampla.

O LRM apresenta avanços em relação ao FRBR, tornando-o mais alinhado com os requisitos atuais de dados abertos e aplicações semânticas. Essa

característica do modelo pode facilitar o mapeamento com o vocabulário do *Schema.org*.

5 CONSIDERAÇÕES FINAIS

O LRM oferece diretrizes para uma estrutura conceitual consistente, proporcionando contexto aos dados bibliográficos e permitindo o estabelecimento de relacionamentos entre eles. Além de promover essas conexões, é fundamental que os dados sejam estruturados de acordo com formatos e vocabulários recomendados para a *Web*, visando à abertura dos catálogos e a integração com outros dados. Dessa forma, os recursos poderão ser encontrados pelos usuários com maior praticidade.

Pondera-se que a modelagem dos catálogos pautada em um modelo conceitual como o LRM, em consonância com vocabulários semânticos como o *Schema.org*, pode estabelecer uma estrutura abrangente e flexível para a representação e compartilhamento dos dados dos catálogos na *Web*. Assim, enquanto o LRM promove uma estrutura conceitual consistente, o *Schema.org* pode possibilitar a integração e a descoberta desses dados pelos mecanismos de busca.

Os resultados observados em relação aos trabalhos da OCLC demonstram a eficácia da aplicação de tecnologias semânticas no contexto das bibliotecas. Portanto, esses exemplos podem inspirar e orientar a modernização dos serviços e do tratamento de informações nas bibliotecas brasileiras, destacando a importância de adotar abordagens inovadoras para aprimorar a conexão, a abertura e a descoberta dos registros na *Web*.

Desse modo, este artigo apresentou uma visão inicial sobre os temas LRM e *Schema.org*. As investigações realizadas fornecem uma base para estudos futuros, permitindo o aprofundamento sobre como os metadados do *Schema.org* podem ser efetivamente integrados e mapeados segundo as orientações do LRM.

REFERÊNCIAS

AMAZON. **CreativeWork Intents**. 2023. Disponível em:

<https://developer.amazon.com/ja-JP/docs/alexa/custom-skills/creativework-intents.html>. Acesso em: 11 jul. 2024.

ALRASHED, Tarfah. *et al.* Dataset or not? A study on the veracity of semantic markup for dataset pages. *In: INTERNATIONAL SEMANTIC WEB CONFERENCE, 20., 2021, [s.l.]. Proceedings [...]. [S.l.]: IBM, 2021. p. 2-17.* Disponível em: <http://people.csail.mit.edu/tarfah/papers/dataset.pdf>. Acesso em: 12 dez. 2024.

ARAKAKI, Ana Carolina Simionato. O Modelo IFLA Library Reference Model e o Linked Data. **Informação & Informação**, Londrina, v. 25, n. 3, p. 163–186, 2020. Disponível em: <http://dx.doi.org/10.5433/1981-8920.2020v25n3p163>. Acesso em: 07 ago. 2024.

BELZ, Olexandra. Use of Schema.org micro-markup in e-commerce projects. **Three Seas Economic Journal**, Riga, v. 3, n. 4, 2022. Disponível em: <http://baltijapublishing.lv/index.php/threeseas/article/view/1964>. Acesso em: 12 dez. 2024.

GALEFFI, Agnese *et al.* **Declaração dos Princípios Internacionais de Catalogação (PIC)**. Tradução de Marcelo Votto Texeira. 2016. Disponível em: <https://repository.ifla.org/handle/123456789/89>. Acesso em: 08 jul. 2024.

GARCÍA, Roberto; GIL, Rosa. Social media copyright management using Semantic Web and Blockchain. *In: INTERNATIONAL CONFERENCE ON INFORMATION INTEGRATION AND WEB-BASED APPLICATIONS & SERVICES, 21., 2019, Munich. Proceedings [...]. Munich: Association for Computing Machinery, 2019. p. 339-343.* Disponível em: <https://dl.acm.org/doi/10.1145/3366030.3366128>. DOI 10.1145/3366030.3366128. Acesso em: 12 dez. 2024.

GIL, Antonio Carlos. **Métodos e técnicas de pesquisa social**. 6. ed. São Paulo: Atlas, 2008.

GODBY, Carol Jean. A division of labor: the role of Schema.org in a Semantic Web model of library resources. *In: GODBY, Carol Jean. Linked data for cultural Heritage*. Chicago: Amer Library Assn Editions, 2016. 152 p. Disponível em: <https://www.oclc.org/content/dam/research/publications/2017/godby-division-of-labor-2017.pdf>. Acesso em: 21 jan. 2024.

GOOGLE. **How Google's Knowledge Graph works**. 2024. Disponível em: <https://support.google.com/knowledgepanel/answer/9787176?hl=en>. Acesso em: 20 jul. 2024.

GOOGLE SEARCH CENTRAL. **Book actions (Book) structured data**. 2024a. Disponível em: <https://developers.google.com/search/docs/appearance/structured-data/book?hl=pt-br#create-your-feed>. Acesso em: 12 jul. 2024.

GOOGLE SEARCH CENTRAL. **Introduction to structured data markup in google search.** 2024b. Disponível

em:<https://developers.google.com/search/docs/appearance/structured-data/intro-structured-data>. Acesso em: 14 out. 2024.

HITZLER, Pascal. A review of the semantic web field. **Communications of the ACM**, Nova York, v. 64, n. 2, 2021. Disponível em:

<https://dl.acm.org/doi/abs/10.1145/3397512>. Acesso em: 09 ago. 2024.

JONES, Shawn, M. *et al.* It's all about the cards: sharing on social media probably encouraged HTML metadata growth. **arXiv**, Nova York, 2021. Disponível em:

<https://arxiv.org/abs/2104.04116>. Acesso em: 12 dez. 2024.

KEJRIWAL, Mayank. *et al.* Empirical best practices on using product-specific Schema.org. *In: AAI CONFERENCE ON ARTIFICIAL INTELLIGENCE*, 35., 2021, Palo Alto. **Proceedings [...]**. Palo Alto: Association for the Advancement of Artificial Intelligence, 2021. p. 15452-15457. Disponível em:

<https://ojs.aaai.org/index.php/AAAI/article/view/17816>. Acesso em: 12 dez. 2024.

META. **Tags de microdados.** 2024. Disponível em:

<https://developers.facebook.com/docs/marketing-api/catalog/guides/microdata-tags>. Acesso em: 11 jul. 2024.

NATIONAL LIBRARY BOARD. **About us.** 2024. Disponível

em:<https://www.nlb.gov.sg/main/about-us>. Acesso em: 11 jul. 2024.

PAYNE, Karen; VERHEY, Chantelle. Schema.org for research data managers: a primer. **International Journal of Big Data Management**, Genebra, v. 2, n. 2, 2022. Disponível em:

<https://www.inderscienceonline.com/doi/10.1504/IJBDM.2022.128449>. Acesso em: 12 dez. 2024.

PEETERS, Ralph. *et al.* Using Schema.org annotations for training and maintaining product matchers. *In: INTERNATIONAL CONFERENCE ON WEB INTELLIGENCE*, 10., 2020, Biarritz. **Proceedings [...]**. Biarritz: Association for Computing Machinery, 2020. p. 195-204. Disponível em: <https://dl.acm.org/doi/10.1145/3405962.3405964>. Acesso em: 2012 dez.2024.

PINTEREST. **Createrich Pins.** 2024. Disponível em:

<https://help.pinterest.com/en/business/article/rich-pins>. Acesso em: 11 jul. 2024.

RIVA, Pat; LE BOEUF, Patrick; ŽUMER, Maja. **IFLA Library Reference Model: a conceptual model for bibliographic information.** Netherlands: International Federation of Library Associations and Institutions, 2017. Disponível em:

https://www.ifla.org/files/assets/cataloguing/frbr-lrm/ifla-lrm-august-2017_rev201712.pdf. Acesso em: 08 jul. 2024.

SERRA, Liliane Giust; SANTARÉM SEGUNDO. Dos silos de dados à web dos dados: bibliotecas e o linked data. **Informação & Informação**, Londrina, v. 26, n. 2, p. 625-645, abr./jun. 2021. Disponível em:

<https://ojs.uel.br/revistas/uel/index.php/informacao/article/view/41603>. Acesso: 08 out. 2024.

SCHEMA.ORG. **Schema.org**. 2024a. Disponível em: <https://schema.org>. Acesso em 10 jul. 2024.

SCHEMA.ORG. **Data model**. 2024b. Disponível em: <https://schema.org/docs/datamodel.html>. Acesso em: 14 jul. 2024.

SCHEMA.ORG. **Full hierarchy**. 2021c. Disponível: <https://Schema.org/docs/full.html>. Acesso em: 17 ago. 2024.

SILVA, L. C. da. **Publicação de dados de pesquisa científica**: proposta de estruturação semântica de cadernos abertos de pesquisa frente às dimensões da e-Science. 2020. 243 f. Tese (Doutorado em Ciência da Informação) - Faculdade de Filosofia e Ciências, Universidade Estadual Paulista Júlio de Mesquita Filho, Marília, 2020. Disponível em: <https://repositorio.unesp.br/handle/11449/194341>. Acesso em: 12 dez. 2024.

TOMOYOSE, Kazumi; SANTOS, Amanda Azevedo dos; ARAKAKI, Ana Carolina Simionato. Schema.org para recuperação da informação em redes sociais. *In*: BRAGATO, T. H. B.; TOGNOLI, N. B. (org.). **Organização do conhecimento responsável**: promovendo sociedades democráticas e inclusivas. Belém: ISKO-Brasil; UFPA, 2019, v. 5, p. 176-182. Disponível em: <https://brapci.inf.br/index.php/res/v/123290>. Acesso em: 12 dez. 2024.

WALLIS, Richard. Follow me to the library! bibliographic data in a discovery driven world. **Italian Journal of Library, Archives and Information Science**, Macerata, v. 13, n. 1, p. 38-44, jan. 2022. Disponível em: <https://www.jlis.it/index.php/jlis/article/view/417>. Acesso em: 12 dez. 2024.

WU, Mingfanget al. An Analysis of Crosswalks from Research Data Schemas to Schema.org. **Data Intelligence**, Pequim, v. 5, n. 1, 2023. Disponível em: <https://direct.mit.edu/dint/article/5/1/100/113281/An-Analysis-of-Crosswalks-from-Research-Data>. Acesso em: 12 dez. 2024.

Notas e créditos do artigo

- **Reconhecimentos:** Os autores gostariam de agradecer ao Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT).
- **Financiamento:** Este estudo foi financiado pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) e pelo Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT).
- **Conflitos de interesse:** Não se aplica.
- **Aprovação ética:** Não se aplica.
- **Disponibilidade de dados e materiais:** Não se aplica.
- **Manuscrito publicado como *preprint*:** o manuscrito foi originalmente publicado como trabalho completo no Seminário Nacional de Catalogação e Tecnologia (SNCat), em 2024. Posteriormente, passou por nova avaliação *double-blind peer review*, além de receber ajustes e atualizações de conteúdo.
- **Contribuições dos autores:**

Contribuição	MACHADO, D. O. F.	ARAKAKI, A. C. S.
Concepção do estudo	X	
Conceitualização	X	X
Metodologia	X	X
Coleta de dados / investigação	X	
Curadoria de dados	X	
Análise dos dados	X	
Discussão dos resultados	X	
Visualização (gráficos, tabelas e outros)	X	
Rascunho original	X	
Revisão e edição final	X	X
Supervisão e administração	X	X
Aquisição de financiamento		X

- **Licença de uso**

Os autores cedem ao **Ciência da Informação Express - CIExpress** direitos exclusivos de primeira publicação, como trabalho simultaneamente licenciado sob a Licença *Creative Commons Attribution (CC BY) 4.0 International*. Esta licença permite que terceiros remixem, adaptem e criem a partir do trabalho publicado, atribuindo o devido crédito de autoria e publicação inicial neste periódico.

• **Publicador**

Universidade Federal de Lavras (UFLA).

As ideias expressas neste artigo são de responsabilidade de sua autoria, não representando, necessariamente, a opinião dos editores ou da universidade.

Editor do canal de comunicação e divulgação científica Ciência da Informação Express

Nivaldo Calixto Ribeiro, Universidade Federal de Lavras (UFLA).

• **Histórico**

Recebido em: 14/11/2024

Aceito em: 21/12/2024

Publicado em: 20/01/2025