



Encontros Bibli: revista eletrônica de biblioteconomia e ciência da informação

ISSN: 1518-2924

Programa de Pós-graduação em Ciência da Informação -
Universidade Federal de Santa Catarina

JESUS, Ananda Fernanda de; CASTRO, Fabiano
Ferreira de; RAMALHO, Rogério Aparecido Sá
O PAPEL DAS BIBLIOTECAS NO LINKED DATA

Encontros Bibli: revista eletrônica de biblioteconomia e
ciência da informação, vol. 26, e75909, 2021, Janeiro-Abril
Programa de Pós-graduação em Ciência da Informação - Universidade Federal de Santa Catarina

DOI: <https://doi.org/10.5007/1518-2924.2021.e75909>

Disponível em: <https://www.redalyc.org/articulo.oa?id=14768130005>

- Como citar este artigo
- Número completo
- Mais informações do artigo
- Site da revista em redalyc.org

UFRJ
redalyc.org

Sistema de Informação Científica Redalyc
Rede de Revistas Científicas da América Latina e do Caribe, Espanha e Portugal
Sem fins lucrativos acadêmica projeto, desenvolvido no âmbito da iniciativa
acesso aberto



Encontros Bibli

O PAPEL DAS BIBLIOTECAS NO *LINKED DATA*

The role of libraries in *Linked Data*

Ananda Fernanda de JESUS

Mestranda em Ciência da Informação

Universidade Federal de São Carlos - Programa de Pós-Graduação em Ciência da Informação, São Carlos, Brasil.

anandafdj@gmail.com

<https://orcid.org/0000-0001-7873-6040> 

Fabiano Ferreira de CASTRO

Doutor em Ciência da Informação

Universidade Federal de São Carlos - Programa de Pós-Graduação em Ciência da Informação, São Carlos, Brasil.

fabianocastro@ufscar.br

<https://orcid.org/0000-0002-8712-2654> 


Rogério Aparecido Sá RAMALHO

Doutor em Ciência da Informação

Universidade Federal de São Carlos - Programa de Pós-Graduação em Ciência da Informação, São Carlos, Brasil.

ramalho@ufscar.br

<https://orcid.org/0000-0002-8491-3514> 

A lista completa com informações dos autores está no final do artigo 

RESUMO

Objetivo: Buscou-se um maior entendimento do papel das bibliotecas no *Linked Data*, visando identificar: 1) como ocorre a representação de recursos informacionais no *Linked Data*; 2) como as bibliotecas podem contribuir com o *Linked Data*; 3) como as bibliotecas podem se beneficiar dos dados publicados pelas demais fontes de dados.

Métodos: Foi realizada uma análise exploratória da temática, pautada em levantamento bibliográfico na Base de Dados em Ciência da Informação (BRAPCI) e no Portal de Periódicos da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES). Além disso, foram analisados os documentos oficiais do *World Wide Web Consortium (W3C)*, responsável por acompanhar o desenvolvimento da *Web* e do *Linked Data*.

Resultado: Identificou-se que o processo de cooperação entre as bibliotecas e as fontes de informação externas ao domínio bibliográfico pode ocorrer a partir de duas vertentes, a de publicadora e a de consumidora de dados. Cada uma dessas formas de atuação pode ser realizada de maneira independente, exigem tomadas de decisão distintas e implicam em diferentes vantagens.

Conclusões: Conclui-se que as bibliotecas devem adotar as recomendações do W3C para potencializar o reuso de seus dados conectados. As instituições do domínio bibliográfico precisam se mobilizar para elaborar instrumentos e critérios que possam embasar sua atuação enquanto consumidora de dados.

PALAVRAS-CHAVE: Dados bibliográficos. *Linked data*. Representação de recursos. *Web*

ABSTRACT

Objective: A greater understanding of the role of libraries in *Linked Data* was sought, aiming to identify: 1) how the representation of information resources in *Linked Data* occurs; 2) how libraries can contribute to *Linked Data*; 3) how libraries can benefit from data published by other data sources.

Methods: An exploratory analysis of the theme was carried out, based on a bibliographic survey in the Database in Information Science (BRAPCI) and in the Periodical Portal of the Coordination for the Improvement of Higher Education Personnel (CAPES). In addition, the official documents of the *World Wide Web Consortium (W3C)*, responsible for monitoring the development of the *Web* and *Linked Data*, were analyzed.

Result: It was identified that the cooperation process between libraries and sources of information external to the bibliographic domain can occur from two aspects, that of publisher and that of data consumer. Each of these forms of action can be carried out independently, require different decision making and have different advantages.

Conclusions: It is concluded that libraries should adopt the recommendations of the W3C to enhance the reuse of their connected data. It is concluded that the institutions of the bibliographic domain need to be mobilized to elaborate instruments and criteria that can base their performance as a data consumer.

KEYWORDS: Bibliographic data. *Linked data*. Resource representation. *Web*

1 INTRODUÇÃO

A busca por formas mais eficientes de organizar os recursos informacionais levou ao surgimento das primeiras bibliotecas, que evoluíram para acompanhar as mudanças causadas pelo desenvolvimento das tecnologias (SMIT, 2012). Essa busca tornou-se mais urgente no contexto atual, caracterizado pela chamada explosão da informação, onde “[...] o termo explosão de informação é uma metáfora (como é a ‘explosão populacional’) porque nada realmente explodiu, mas apenas cresceu a uma taxa alta, mesmo que exponencialmente”. (SARACEVIC, 2009, p. 1, tradução nossa).

O desenvolvimento da *Web* foi responsável por acelerar a velocidade com que se compartilham os recursos informacionais, entretanto não foram desenvolvidos e posteriormente adotados instrumentos de padronização para a representação dos conteúdos disponibilizados nesse ambiente, causando problemas no processo de recuperação.

Os metadados, “dados que descrevem outros dados em um sistema de informação, com o intuito de identificar de forma única uma entidade (recurso informacional) para posterior recuperação” (ALVES, 2010, p. 47), são cruciais para permitir a recuperação das informações no ambiente *Web*, pois padronizam a representação dos recursos. Com o objetivo de ampliar a utilização e a padronização desses metadados foi proposto um conjunto de práticas denominado *Linked Data*, que consiste na proposta de interligação de dados representados com base em um formato padrão, partindo da utilização de *links* que permitem relacionar dados advindos de diversas fontes.

As bibliotecas são fornecedoras importantes de metadados, devido à sua tradição na utilização de formatos, de códigos e de padrões que garantem a consistência e a veracidade dos seus dados. Os mesmos formatos e padrões que têm garantido ao longo dos anos a consistência dos dados produzidos pelas bibliotecas faz com que eles não sejam recuperados por meio dos buscadores da *Web*, além disso, as novas estruturas de armazenamento, os tipos documentais e os meios de busca e a obtenção de informação advindos da *Web* alteraram o comportamento e as necessidades dos usuários de informação.

Ampliar a utilização do *Linked Data* no contexto das bibliotecas demonstra ser uma forma eficiente de adequar os dados bibliográficos ao cenário informacional e, ao mesmo tempo, de contribuir para melhorar a recuperação da informação no ambiente *Web*.

Arakaki (2016) aponta que muitas bibliotecas no exterior já estão se mobilizando para adequar seus dados ao *Linked Data*, ainda assim, esse processo tem sido lento e cauteloso. Muitas vezes, como apontado por Castro e Jesus (2018), a comunidade se vê reticente em adotar as práticas do *Linked Data*, principalmente por se preocuparem com a heterogeneidade dos dados disponibilizados no ambiente *Web*.

Nesse sentido, entende-se que uma maior compreensão sobre como são representados os recursos informacionais no *Linked Data* e um maior esclarecimento a respeito das formas de atuação das bibliotecas nesse contexto, ajudaria a diminuir o receio da comunidade a respeito da adoção do *Linked Data*. Partindo da identificação dessa necessidade questiona-se: como ocorre a cooperação entre as bibliotecas e demais fontes de informação disponíveis no contexto do *Linked Data*?

Objetivou-se um maior entendimento do papel das bibliotecas no *Linked Data*, visando identificar: 1) como ocorre a representação de recursos informacionais no *Linked Data*; 2) como as bibliotecas podem contribuir com o *Linked Data*; e 3) como as bibliotecas podem se beneficiar dos dados publicados pelas demais fontes de dados.

A pesquisa é caracterizada como exploratória e descritiva, sendo obtidos resultados qualitativos, a partir da análise da literatura científica, a qual possibilitou a construção do *corpus* teórico desse estudo. Como método foi realizada uma análise exploratória sobre o papel das bibliotecas no *Linked Data*, pautada em recursos bibliográficos e documentais. As análises exploratórias não são pesquisas exaustivas, tendo caráter mais abrangente. Para o levantamento do material bibliográfico, utilizou-se a Base de Dados em Ciência da Informação (BRAPCI) e o Portal de Periódicos da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES). A busca foi realizada utilizando os termos “*Linked Data*”, “dados bibliográficos”, “catálogo bibliográfico”, “catalogação” e respectivas versões em português, inglês e espanhol. Foram analisados os artigos que tratam das discussões a respeito de como ocorre a representação de recursos informacionais no *Linked Data* e os referenciais teóricos dos artigos recuperados. Não foi aplicada nenhuma restrição geográfica ou temporal.

Na etapa de levantamento documental foram consideradas as informações contidas nos documentos oficiais do *World Wide Web Consortium (W3C)*, consórcio responsável por acompanhar o desenvolvimento da *Web* e pela elaboração das recomendações do *Linked Data*. Os documentos foram selecionados a partir dos seguintes critérios: (1) documentos que se aprofundassem em cada uma das quatro recomendações do *Linked Data*; (2) documentos que tratassem da relação das melhores

práticas com o ambiente das bibliotecas; e (3) documentos que pudessem potencializar a utilização e a publicação de dados ligados. Com base nos resultados da análise exploratória discutiu-se como ocorre a representação dos recursos informacionais no *Linked Data*, subsidiada pelo detalhamento de cada recomendação e da apresentação de um conjunto de práticas que pode potencializar a publicação de dados conectados (seção 2). Apresenta-se também, como as bibliotecas podem contribuir e se beneficiar do *Linked Data*, partindo de um breve histórico da evolução dos catálogos e das vantagens da adoção do *Linked Data* nesses ambientes informacionais (seção 3).

2 BREVE EVOLUÇÃO DA REPRESENTAÇÃO DE RECURSOS INFORMACIONAIS

O desenvolvimento da imprensa por Gutenberg no século XV levou a popularização dos recursos informacionais, a informação registrada passou a poder ser replicada com perfeição e velocidade, tornando-se mais acessível. Os documentos impressos são elaborados para possibilitar a leitura linear de seus conteúdos. Críticas às limitações provocadas pela linearidade dos documentos impressos e dos próprios sistemas de organização a que eram condicionados esses documentos são a base das inquietações do clássico artigo de Bush (1945), intitulado *As we may think*. A respeito dessa linearidade o autor aponta que:

A mente humana não funciona desta forma, ela opera por meio de associações. Quando um elemento está ao seu alcance, salta instantaneamente para o seguinte, que é sugerido pela associação de pensamentos segundo uma intrincada rede de atalhos contida nas células do cérebro. (BUSH, 1945, p. 8, tradução nossa).

O ponto central do artigo de Bush é a busca por formas mais orgânicas de disponibilização, de organização e de recuperação da informação, mais aproximadas daquela que ocorre na mente humana, ou seja, de maneira não linear por meio de interligação entre os recursos.

As potencialidades desse tipo de ligação dentro de um mesmo documento e entre documentos distintos foram ampliadas graças aos avanços na área da computação. Nesse contexto tecnológico destaca-se o conceito de hipertexto, o termo foi cunhado por Theodore Nelson e se relaciona “à ideia de leitura/ escrita não-linear em sistemas informatizados. Durante suas pesquisas, Nelson descreveu muitas das ideias implantadas nos sistemas hipertexto atuais”. (DIAS, 2019, p. 272).

Desde a criação do hipertexto, diversos projetos computacionais foram desenvolvidos com base nesse conceito, dentre os quais destaca-se a criação do ambiente *Web*, proposto para permitir o compartilhamento de recursos multimídia e a interligação desses recursos partindo da utilização de hipertexto.

O documento *Web* é construído sobre um pequeno conjunto de padrões simples: Identificadores Uniformes de Recursos (URIs) como mecanismo de identificação globalmente único, o Protocolo de Transferência de Hipertexto (HTTP) como mecanismo de acesso universal e a Linguagem de Marcação de Hipertexto (*HTML*) como formato de conteúdo amplamente utilizado. Além disso, a *Web* se baseia na ideia de configurar hiperlinks entre documentos da *Web* que podem residir em servidores *Web* diferentes. (HEATH; BIZER, 2011, não paginado, tradução nossa).

A base para a inserção de recursos na *Web* é a linguagem de marcação *Hypertext Markup Language (HTML)*. Ela foi desenvolvida para descrever as estruturas das páginas criadas na *Web*, permitindo a inserção de informações multimídia e a criação de *hyperlinks*, ou seja, a ligação entre os recursos contidos em uma mesma página ou em páginas diversas. Por meio do *HTML* são criados *links*, “um link é uma conexão de um recurso da *Web* para outro. Embora seja um conceito simples, o link tem sido uma das principais forças impulsionadoras do sucesso da *Web*”. (W3C, 2018, não paginado, tradução nossa). A ligação de recursos informacionais em *HTML* acontece da seguinte maneira:

Um link tem duas extremidades - chamadas âncoras - e uma direção. O link começa na âncora “source” e aponta para a âncora “destination”, que pode ser qualquer recurso da *Web* (por exemplo, uma imagem, um videoclipe, uma frase de áudio, um programa, um documento *HTML*, um elemento dentro de um *HTML* documento, etc.). (W3C, 2018, não paginado, tradução nossa).

A *HTML* foi estruturada para ser de fácil compreensão e com foco na apresentação dos conteúdos,

Contudo, sua concepção baseada em marcações fixas trouxe limitações para a representação de informações uma vez que, sendo uma linguagem de marcação voltada, principalmente, para a apresentação dos documentos em ambientes digitais, *HTML* apenas define a forma como a informação é exibida não se preocupando com o significado da palavra. (RAMALHO; MARTINS; SOUZA, 2017, p. 25).

As limitações relacionadas à representação das informações contidas na *Web* têm consequência direta na recuperação dessas informações. Sendo os metadados indispensáveis para a representação dos recursos na *Web*, passou-se a buscar formas mais eficientes de promover seu uso e reuso.

Heath e Bizer (2011) apontam que, como forma de minimizar a dificuldade de coleta e reuso dos metadados foram criados microformatos – adaptações em padrões existentes para a solução de problemas específicos na *Web* – e *Application Programming Interface (APIs)* – conjunto de rotinas e padrões que permitem que aplicativos possam se comunicar e realizar atividades de maneira automatizada sem a intervenção dos usuários.

Mas essas iniciativas são paliativas, pois os microformatos permitem “[...] representar dados sobre um pequeno conjunto de tipos diferentes de entidades” (HEATH; BIZER, 2011, não paginado, tradução nossa), enquanto que “[...] a existência de uma API especializada para cada conjunto de dados cria um cenário em que um esforço significativo é necessário para integrar cada conjunto de dados original em um aplicativo”. (HEATH; BIZER, 2011, não paginado, tradução nossa).

Uma das alternativas para ampliar o uso e o reuso de metadados na *Web* é o *Linked Data*, proposto pelo W3C e composto por um conjunto de práticas para promover a descrição e a interligação de metadados. De acordo com Berners-Lee (2006, não paginado, tradução nossa), o *Linked Data* consiste nas seguintes práticas:

1. Use URIs como nomes para as coisas.
2. Use HTTP URIs, para que as pessoas possam procurar esses nomes.
3. Quando alguém procura um URI, forneça informações úteis, usando os padrões (*RDF*, *SPARQL*).
4. Inclua links para outros URIs, para que eles possam descobrir mais coisas.

Os *Uniform Resource Identifier (URIs)* são identificadores que servem para representar documentos, locais, endereços da *Web*, figuras ou até mesmo pessoas. O “[...] URI é um padrão conjunto que abarca os conceitos de *Uniform Resource Locator* (URL) e do *Uniform Resource Name* (URN), de modo que pode ser representado por qualquer um destes, ou por ambos”. (RAMALHO, 2006, p. 32). Enquanto o *URL* é utilizado para marcar a localização na *Web*, o *URN* é utilizado para especificar o nome do recurso.

O *Hypertext Transfer Protocol (HTTP)*, de maneira geral, é um protocolo de solicitação/resposta de uso padrão, que permite acesso às informações contidas no ambiente *Web*, realizando a conexão entre dois computadores conectados à internet. (W3C, 2018, não paginado). “Assim, o segundo princípio *Linked Data* defende o uso de URIs *HTTP* para identificar objetos e conceitos abstratos, permitindo que esses URIs sejam desreferenciados, ou seja, pesquisados” (HEATH; BIZER, 2011, não paginado, tradução nossa).

O *Resource Description Framework (RDF)* é um modelo proposto em 1999 com o objetivo de promover interoperabilidade e padronização dos recursos hospedados na *Web*.

[...] o *RDF* usa uma terminologia específica para falar sobre as várias partes das declarações. Especificamente, a parte que identifica o assunto da declaração é chamada de assunto. A parte que identifica a propriedade ou característica do assunto que a instrução específica (criador, data de criação ou idioma nesses exemplos) é chamada de predicado e a parte que identifica o valor dessa propriedade é chamada de objeto. (W3C, 2004, não paginado, tradução nossa).

Segundo o W3C (2004), o *RDF* é um modelo utilizado quando as relações entre os dados precisam ser descritas de maneira formal, ou seja, legível por máquinas. Para que ele atenda a esse objetivo, os seguintes pontos precisam ser levados em consideração: (1) As coisas descritas (assunto) e as propriedades (predicado) precisam ser representadas utilizando *URIs*; (2) O valor das declarações (objetos) pode ser descrito usando um *URI* ou um *literal* (números, nomes, palavras, frases em linguagem natural); (3) Para cada nova propriedade do assunto é criada uma nova tripla; (4) Para ser legível por máquinas, ele precisa do suporte de uma linguagem computacional - um formato de serialização; e (5) É necessário criar e utilizar vocabulários controlados para descrever as propriedades das coisas descritas.

A linguagem computacional utilizada para a serialização e o vocabulário utilizado para descrever as propriedades dos assuntos podem variar de acordo com as necessidades da instituição. Os vocabulários são criados para atender a domínios específicos, o W3C recomenda que sempre que possível sejam adotados vocabulários existentes, sendo possível aplicar mais de um vocabulário simultaneamente. Quando as propriedades necessárias não são encontradas em nenhum vocabulário, um novo pode ser criado, e é recomendado que esse seja amplamente divulgado para que possa ser reutilizado por outras entidades pertencentes ao mesmo domínio.

Também é possível representar as triplas *RDF* no formato de grafos, o que permite uma melhor compreensão para usuários humanos “por convenção, os recursos são representados como elipses, os valores literais como retângulos e os predicados utilizando arcos direcionados do recurso (sujeito) para o valor (objeto)” (RAMALHO; MARTINS; SOUZA, 2017, p. 29). A figura 1 representa uma declaração em linguagem natural e a mesma declaração em formato de grafo *RDF*.

Figura 1 – Exemplo de conversão de linguagem natural para *RDF*



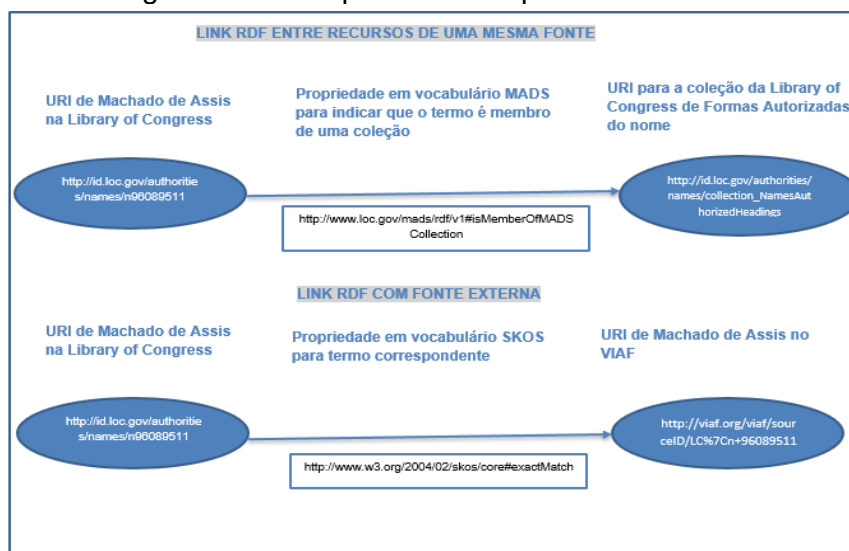
Fonte: Autores (2020)

Na figura 1 pode-se observar o exemplo de uma declaração envolvendo a autora Clarice Lispector e um de seus pseudônimos, Tereza Quadros, onde a intenção é explicitar a relação existente entre as duas formas de nome adotadas pela autora. Em linguagem natural, uma das possibilidades de indicação dessa relação é por meio da frase “Clarisse Lispector tem como termo alternativo Tereza Quadros”. Para representar a mesma relação em *RDF* é necessário o uso de um *URI* para o assunto da frase e uma propriedade em um vocabulário controlado que seja equivalente a essa relação.

No exemplo apresentado na figura 1 utilizou-se o *URI* para Clarice Lispector estabelecido pelo *Virtual International Authority File (VIAF)*, um consórcio criado com a finalidade de agrupar os dados de autoridade de diversas bibliotecas nacionais sob um único identificador representado por um *URI*. Para representar o predicado da frase utilizou-se como exemplo, a propriedade *altLabel*, que indica um termo alternativo, estabelecida pelo vocabulário *Simple Knowledge Organization System (SKOS)*, criado com a finalidade de permitir “[...] a representação de Sistemas de Organização do Conhecimento em ambientes digitais”. (RAMALHO; SOUSA, 2019, p. 126).

Existem dois tipos principais de declarações em *RDF*, aqueles que descrevem as propriedades de um recurso, nas quais o objeto é representado em linguagem natural, chamados de triplas literais, e os que representam as ligações existentes entre dois recursos, onde o objeto é outro *URI*, chamados *Links RDF* (HEATH; BIZER, 2011). A Figura 2 apresenta exemplo de *links RDF* que relacionam dados de uma mesma fonte e *links* que relacionam dados de fontes externas, quando os objetos são outros *URIs* também convencionou-se representá-los utilizando elipses.

Figura 2 – Exemplo dos dois tipos de Links *RDF*



Fonte: Autores (2020)

De acordo com a Figura 2, o Grafo 1 apresenta um exemplo de *link* entre dados contidos no catálogo de autoridades da *Library of Congress* (LC). Esse grafo poderia ser representado em linguagem natural pela frase “A autoridade para Machado de Assis faz parte do catálogo de autoridades da LC”. Para representá-la em *RDF* foram utilizados os *URIs* da LC criado para Machado de Assis e para o seu catálogo de autoridades. Também foi utilizada uma propriedade criada para representar relações de pertencimento do vocabulário *Metadata Authority Description Schema* (MADS), elaborado para representar relações existentes em vocabulário controlados de instituições voltadas para o patrimônio cultural.

No Grafo 2, da figura 2, é possível observar um exemplo de *link* entre fontes distintas, nele a frase “a autoridade Machado de Assis da LC corresponde à autoridade Machado de Assis do VIAF”, em *RDF* é representado pelo *URI* de Machado de Assis, na LC e no VIAF e pela propriedade criada para termo corresponde no SKOS.

Nesse sentido, observa-se que tal como o *hiperlink* que interliga recursos em *HTML* contidos em uma mesma página ou em páginas distintas, os *links* em *RDF* podem fornecer ligações entre dados de uma mesma fonte ou com dados de fontes externas. Observa-se ainda, que *URIs* de fontes distintas podem ser usados para representar uma mesma entidade, e que vocabulários distintos podem ser utilizados em uma mesma descrição, de acordo com o que se deseja representar.

Como forma de auxiliar no processo de publicação de dados na *Web*, o W3C (2017) disponibilizou um guia contendo 35 (trinta e cinco) recomendações formuladas para garantir que os dados publicados na *Web* sejam recuperados e passíveis de

reutilização. O Quadro 1 apresenta a relação entre as recomendações do W3C (2017), sua descrição e, sempre que pertinente, a contextualiza ao domínio das bibliotecas.

Quadro 1 – Recomendações para a publicação e utilização de dados na Web

Recomendação	Descrição
1- Fornecer Metadados	As bibliotecas já fornecem metadados como parte da sua atividade cotidiana, o diferencial é que esses metadados devem ser compartilhados em formato legível por humanos e por computadores, sendo possíveis duas abordagens: (1) fornecer os dados como parte de uma página da Web em HTML (2) fornecer os metadados como um arquivo de texto separado.
2- Fornecer metadados descritivos	Essa recomendação diz respeito a criar metadados para descrever o próprio conjunto de dados, como forma de explicitar a natureza e facilitar a descoberta desses dados. Os metadados desejáveis são: título, descrição, palavra-chave, data de publicação, entidade responsável, dados de contato, cobertura temporal, cobertura espacial e data da última modificação.
3- Fornecer metadados estruturais	Diz respeito a fornecer informações sobre a estrutura do conjunto de dados, essa prática é particularmente importante para o processamento automatizado dos dados, para que os usuários computadores possam identificar o esquema e a estrutura dos dados realizando atividades de interoperabilidade.
4- Fornecer informações de licença de dados	Diz respeito a selecionar uma licença que permita aos futuros consumidores saber de que maneira o conjunto de dados pode ser utilizado e as possíveis restrições a esse reuso. Segundo Rautenberg et al. (2018, p. 35), as seguintes questões devem nortear a escolha da licença adequada: <ul style="list-style-type: none"> • Serão permitidas alterações nos dados? • Os dados podem ser utilizados para fim comercial? • Os direitos autorais anteriores devem ser mantidos? • Os resultados da utilização desses dados devem manter a mesma licença?
5- Fornecer informações de proveniência de dados	A importância de fornecer informações sobre a origem dos dados está relacionada a preocupação com a heterogeneidade da Web, considerando inclusive que muitas vezes a fonte que está publicando determinados dados não é a mesma que os criou. Saber a origem dos dados é o que permite que os consumidores possam confiar na integridade e credibilidade dos mesmos.
6- Fornecer informações de qualidade de dados	Diz respeito a fornecer informações que permitam a seleção de conjuntos de dados a partir do seu nível de qualidade, aumentando as chances de reutilização. Nesse sentido a entidade deve levar em consideração os critérios de qualidade do seu próprio domínio (como os fornecidos pelos códigos e padrões da biblioteconomia) e ainda documentar os problemas de qualidade conhecidos que possam ser relevantes para outros domínios. Para esse fim o W3C (2017) fornece o Vocabulário de Qualidade de Dados (VOCAB-DQV) que define conceitos para avaliar a qualidade.
7- Fornecer um indicador de versão	Os dados publicados podem precisar ser alterados com o tempo, por mudanças tecnológicas, alterações nos instrumentos de padronização ou mesmo porque se encontravam desatualizados. Usar um identificador de versão permite que os usuários localizem as versões mais recentes ou saibam que dados que eles estão utilizando sofreram alterações.

Recomendação	Descrição
8- Fornecer histórico de versões	Essa prática está diretamente relacionada à prática anterior, e diz respeito a manter um histórico que indique o que foi alterado de uma versão para a outra, permitindo aos usuários identificar se essa alteração é ou não pertinente para o seu contexto de uso.
9- Usar URIs persistentes como identificadores de conjuntos de dados	Um ponto importante na avaliação da confiabilidade dos dados é garantir que os <i>URIs</i> sempre façam referência ao mesmo recurso ao longo do tempo. Os <i>URIs</i> devem direcionar para o mesmo conjunto de dados mesmo que esses tenham sido alterados ou caso esses não estejam mais disponíveis remeter a uma informação que indiquem o destino desses dados.
10- Utilizar URIs persistentes de fontes confiáveis no conjunto de dados	Diz respeito a criar <i>Links RDF</i> com os <i>URIs</i> de outras fontes nos conjuntos de dados sempre que isso for possível. Essa recomendação retoma a última das melhores práticas do <i>Linked Data</i> , e é justamente sobre construir ligações com fontes externas.
11- Atribuir URIs a versões e séries de conjuntos de dados	Diz respeito a criar identificadores para novas versões dos dados e ainda sobre a identificação de grupos de dados. Muitos conjuntos de dados se enquadram em séries ou grupos naturais, como por exemplo, os dados de autoridade contidos nos catálogos das bibliotecas.
12- Usar formatos de dados padronizados legíveis por máquina	Os conjuntos de dados se tornam maiores e mais complexos, sendo crucial a utilização de formatos que permitam o seu processamento por computadores. A padronização sempre foi a chave da consistência dos dados no ambiente das bibliotecas e isso se repete no contexto do <i>Linked Data</i> , usar formatos padronizados amplia as possibilidades de reuso dos dados e ainda evita o retrabalho.
13- Usar representações de dados com neutralidade de localidade	Diz respeito à sempre que possível optar por dados que são legíveis por máquina e não específicos a qualquer idioma ou cultura. Mas nem sempre isso é possível, nesse caso a recomendação é que sejam fornecidos metadados que especifiquem as características dos dados.
14- Fornecer dados em vários formatos	Os dados, incluído as declarações em <i>RDF</i> como foi discutido anteriormente, podem ser serializados em formatos diferentes, fornecer os mesmos dados em vários formatos de serialização amplia as possibilidades de uso desses dados, minimiza as possibilidades de erro em processos de conversão.
15- Reutilizar vocabulários, preferencialmente padronizados	Muitos vocabulários podem ser utilizados para representar as propriedades dos recursos, reutilizar vocabulários existentes evita retrabalho, capta e facilita o consenso nos domínios, aumenta a interoperabilidade, evita ambiguidades e ajuda na comparação e no processamento automático de dados e metadados.
16- Escolher o nível de formalização correto	Diz respeito à escolha dos vocabulários utilizados para tornar os dados legíveis por máquina. Vocabulários complexos exigem mais esforço, escolher um vocabulário muito simples pode ser mais prático, mas a sua simplicidade pode não ser suficiente para representar um conjunto de dados complexo. Deve-se sempre buscar um equilíbrio entre a maneira mais simples e eficiente de representar os dados.
17- Fornecer o <i>download</i> em massa	Diz respeito a permitir que os consumidores recuperem o conjunto de dados completo com uma única solicitação. Mesmo que as informações em <i>RDF</i> sejam divididas em pequenas declarações no formato de triplas, geralmente a informação pode ser naturalmente agrupada em um conjunto maior de informações, como todas as triplas que dizem respeito a um autor específico ou mesmo todos os

Recomendação	Descrição
	arquivos de autoridade da biblioteca, nesse caso deve-se permitir que com uma única ação todo o conjunto de declarações seja baixado.
18- Fornecer subconjuntos para conjuntos de dados grandes	Nem sempre o usuário estará interessado no conjunto completo de dados, nesse caso, retomando o exemplo da recomendação 17, é interessante permitir que o usuário possa optar por baixar todas as triplas de um determinado autor ou o arquivo de autoridade completo, dependendo da sua necessidade.
19- Usar a negociação de conteúdo para conectar dados disponíveis em vários formatos	A negociação de conteúdo permite que com um mesmo <i>URL</i> o usuário tenha acesso a todos os formatos em que uma mesma informação está disponível e possa optar por aquela que lhe interessa.
20- Fornecer acesso em tempo real	Essa recomendação é destinada a dados que sofrem alterações constantes como no caso de estatísticas de acesso e disponibilidades de determinados recursos. Para que eles sejam úteis é necessário que eles também sejam disponibilizados em tempo real. A necessidade de fornecer acesso em tempo real precisará ser avaliada caso a caso, considerando as taxas de atualização, o tempo decorrido nas etapas de pós-processamento de dados, a disponibilidade da infraestrutura e as necessidades dos consumidores em potencial.
21- Fornecer dados atualizados	Diz respeito a sempre disponibilizar as versões mais recentes dos dados bem como fornecer informações sobre a frequência com que esses dados costumam ser atualizados.
22- Forneça uma explicação para os dados que não estão disponíveis	Com o uso de URIs permanentes é necessário fornecer ao usuário uma explicação para os dados que não estão mais disponíveis. Para realizar essa atividade essas páginas podem retornar os erros pré-estabelecidos da linguagem <i>HTML</i> , por exemplo, o 303, que indica que a informação esta disponível em outro lugar; 410 que indica que a informação foi definitivamente excluída e o 503, que indica que a mesma está temporariamente indisponível.
23- Disponibilizar dados por meio de uma API	APIs são conjuntos de rotinas e de padrões que permitem que aplicativos possam se comunicar e realizar atividades de maneira automatizada. O uso de APIs em combinação com os conjuntos de dados oferece flexibilidade e capacidade de processamento, podendo, por exemplo, ativar o uso de dados em tempo real, moderar a negociação de conteúdos e a capacidade de <i>download</i> do conjunto completo ou de conjuntos menores. Criar uma API é mais complicado do que simplesmente disponibilizar dados, nesse sentido é possível habilitar e reutilizar APIs existentes.
24- Usar padrões da Web como base de APIs	Indica o uso de um estilo arquitetônico baseado nas tecnologias da própria <i>Web</i> no processo de criação de APIs como forma de aproveitar os pontos fortes da <i>Web</i> .
25- Fornecer documentação completa para sua API	Além da padronização a outra chave para a interoperabilidade e ampla utilização dos dados é a documentação de todos os processos e alterações. Quando as informações sobre as APIs são completas e de fácil compreensão, as chances de que essa seja reutilizada são ampliadas.
26- Evitar quebrar alterações na sua API	Alterar características essenciais da API pode prejudicar os usuários e assim diminuir a confiança desses no conjunto de dados, portanto essas alterações devem ser evitadas sempre que possível, e quando não for possível, elas devem ser comunicadas.

Recomendação	Descrição
27- Preservar identificadores	Reforça a prática 22 sobre o uso de URIs persistente, principalmente sobre informações removidas, não é possível que todos os dados na <i>Web</i> permaneçam disponíveis indefinidamente, mas é fundamental fornecer informações para que o usuário saiba o que aconteceu e ainda que os identificadores não sejam reaproveitados.
28- Avalie a cobertura do conjunto de dados	Para preservar os dados não basta garantir apenas que os arquivos locais estejam salvaguardados, os vocabulários e os formatos necessários para torná-los compreensíveis também precisam ser preservados. Optar por vocabulários e formatos amplamente utilizados diminui os riscos de que os dados deixem de ser compreensíveis.
29- Coletar <i>feedback</i> dos consumidores de dados	Permitir um canal de comunicação entre publicadores e consumidores garante que os dados publicados atendam às necessidades do consumidor e ajuda os publicadores a melhorarem os serviços fornecidos. Essa coleta pode ser feita tanto para usuários humanos, por meios de canais para comentários e contato direto com os publicadores, como para usuários computadores por meio de monitoramento automatizado do uso dos dados.
30- Disponibilizar <i>feedback</i>	Sempre que possível, o <i>feedback</i> também deve estar disponível publicamente, o que cria um ambiente colaborativo, demonstrar aos usuários que suas preocupações foram recebidas e estão sendo processadas, evitado assim envio de relatórios duplicados.
31- Enriquecer dados gerando novos dados	O enriquecimento de dados diz respeito a processos complexos que visam ampliar o alcance dos dados originais. Técnicas para enriquecimento de dados são complexas e não podem ser aplicados a todo tipo de dados, pois podem influenciar negativamente a confiabilidade dos dados. O próprio W3C afirma que “Técnicas para enriquecimento de dados são complexas e vão muito além do escopo deste documento, que pode apenas destacar as possibilidades”. (W3C, 2017, não paginado).
32- Oferecer Apresentações Complementares	Diz respeito criar visualizações interativas ou aplicativos da <i>Web</i> que usem os dados. A disponibilização de utilização prática dos dados pode torná-los mais compreensíveis e reutilizáveis.
33- Fornecer <i>feedback</i> ao Publicador Original	O potencial dos dados ligados está em permitir que esses sejam reutilizados de maneiras distintas, tais como a criação de novas combinações, criação de aplicativos ou visualizações interativas, publicação um novo formato ou idioma. Fornecer aos publicadores <i>feedback</i> sobre como seus dados estão sendo úteis incentiva a manutenção desses dados, a publicação de novos dados e ainda, permite que o publicador faça alterações para facilitar o trabalho do consumidor de dados.
34- Siga os Termos de licença	A licença de dados diz como os dados podem ser utilizados, como os dados resultantes de manipulações devem ser publicados e indica a necessidade de fornecer informações sobre a autoria original dos dados. Caso a licença não esteja explicitada é necessário entrar em contato com os fornecedores originais antes de fazer uso desses dados.
35- Cite a Publicação Original	Fornecer meios para que os usuários identifiquem a fonte original dos dados é um fator importante para que esses possam checar a qualidade e a confiabilidade dos dados. Ao fornecer essa informação o publicador aumenta a confiança do consumidor em seus dados e as chances de reuso dos mesmos.

Fonte: Autores (2020), baseado em W3C (2017).

Embora não tenham sido elaboradas exclusivamente para a aplicação em dados conectados, entende-se que as 35 recomendações apresentadas no quadro 1 podem ser somadas às 4 práticas do *Linked Data*, já que elas podem garantir que os dados conectados sejam mais consistentes, permitir a checagem da qualidade desses dados, garantir que o usuário saiba quem são seus fornecedores e qual a licença de uso desses dados. A adoção dessas recomendações por dados publicados como *Linked Data* aumentaria, portanto, a reutilização de dados, ao fomentar a confiança entre publicadores e consumidores de dados e o potencial de ligação.

3 AS BIBLIOTECAS NO CONTEXTO DO *LINKED DATA*

Desenvolvida no âmbito da Biblioteconomia como forma de estruturar e de padronizar as tentativas até então dispersas de representação das informações contidas nos acervos das bibliotecas “a catalogação é o processo por meio do qual se descreve formalmente um documento ou recurso e se estabelece um variado e variável número de pontos de acesso [...]”. (CATARINO; SOUZA, 2012, p. 84).

A Catalogação Descritiva é responsável por parte do processo de tratamento dos dados e das informações armazenadas nas bibliotecas, ela resulta no catálogo bibliográfico, que permite a identificação dos registros presentes em um acervo e deve agir não só como uma lista ordenada desses conteúdos, mas permitir que o usuário encontre, identifique, selecione e adquira um recurso informacional (DECLARAÇÃO, 2016).

A origem dos catálogos pode ser relacionada ao surgimento da escrita, porém, nesse período predominaram iniciativas que são mais corretamente denominadas como pré-catálogos, pois se assemelhavam a lista de inventário e eram condicionadas às tecnologias disponíveis no período, como as tabuletas de argila e o papiro (ALVES, 2010). Os catálogos estruturados e complexos são um resultado da invenção da imprensa e do aumento exponencial de informações registrada, sendo que os primeiros instrumentos de padronização utilizados na construção do mesmo datam do século XIX.

Um dos primeiros elementos de ligação dos catálogos foram as remissivas ver e ver também, os quais indicam ao usuário outras informações que podem ser relevantes ao seu processo de busca, tais como assuntos, autores e obras correlatas e outras formas de escrita para um determinado assunto, nome ou título. A ligação, desde então,

se estabelece como uma das características importantes para que os catálogos se diferenciasssem de listas de inventário e permitissem que os usuários navegassem entre um registro e outro.

O surgimento dos sistemas *online* e dos microcomputadores ampliaram as possibilidades de acesso aos catálogos, que passaram a ser disponibilizados *online* e com acesso público, os chamados *Online Public Access Catalog (OPAC)*.

O catálogo online é um processo automatizado, no qual uma ferramenta é disposta num banco de dados, que a partir de um determinado servidor armazena e recupera os mais variados tipos de informações em formatos eletrônicos pertencentes à biblioteca, garantindo uma maior velocidade e qualidade de acesso a informação. (QUEIROZ; ARAÚJO, 2013, p. 8).

Com os catálogos *online*, as bibliotecas se beneficiaram das possibilidades oferecidas pelo *HTML* e pelo *hiperlink*, embora as ligações com fontes externas ainda não sejam tão frequentes, a ligação entre informações de um mesmo *OPAC* é o que permite que o usuário navegue no catálogo e obtenha as informações de seu interesse. As próprias remissivas passaram a consistir em *hiperlinks* que direcionam o usuário à próxima informação de interesse.

Com “a disponibilização do Online Public Access Catalogue (OPAC) na *Web*, o alcance dos Catálogos foi ampliado”. (SANTARÉM SEGUNDO; SERRA, 2017, p. 173). Porém, o crescente volume de informação registrada, gerado pelo advento do ambiente *Web* levanta questões sobre como garantir a representação consistente e em tempo hábil, de toda essa informação (ARAKAKI, 2016).

Adotar o *Linked Data* possibilita que os dados das bibliotecas sejam adequados ao cenário informacional, entretanto, essa adequação tem acontecido de maneira cautelosa, sobretudo porque, a comunidade se vê reticente em adotar as práticas do *Linked Data*, o que em partes se deve a uma falta de compreensão sobre como ocorre a ligação no contexto do *Linked Data* e sobre quais papéis as bibliotecas podem exercer nesse contexto (CASTRO; JESUS, 2018).

O interesse em adequar os dados das bibliotecas ao *Linked Data* foi reforçado pelo relatório *Library Linked Data Incubator Group: final report*, resultado de um conjunto de discussões realizadas entre maio de 2010 a agosto de 2011 no âmbito do *W3C*, sobre como incluir as bibliotecas no contexto do *Linked Data*.

O relatório indicou uma série de vantagens, de desafios e de recomendações para ampliar a participação das bibliotecas no *Linked Data* e destaca que “Bibliotecas e instituições de memória em geral estão em uma posição única para fornecer metadados

confiáveis para recursos de importância cultural de longo prazo, como dados na *Web*". (W3C INCUBATOR GROUP, 2011, não paginado, tradução nossa).

O relatório apresenta como principais vantagens da adoção do *Linked Data* pelas bibliotecas, a interoperabilidade de dados com instituições externas ao domínio bibliográfico e a ampliação da presença das bibliotecas na *Web*.

Como vantagens, Agenjo, Hernández e Viedma (2012) e Godby e Smith-Yoshimura (2016) apontam que a estrutura do *Linked Data* permitiria que os dados transpusessem as barreiras linguísticas, sendo possível a ligação entre dados advindos de diversos países.

Han (2016) e Laurence (2013) apontam como vantagem uma maior agilidade no processo de tratamento de novos recursos informacionais. Enquanto Xu et al. (2016), Borie e Van Ballegooie (2014) e Wenz (2013), trazem como vantagem a possibilidade de ligação entre os dados do recurso informacional descrito e os dados de outros recursos que podem ser de interesse para o usuário, como a relação entre uma partitura, sua execução em formato musical e uma apresentação em vídeo.

Castro e Jesus (2018) apontam ainda que dentre as vantagens mais discutidas na literatura, se destacam o aumento da granularidade das representações elaboradas pelas bibliotecas e a obtenção de melhores resultados de busca, permitindo um refinamento mais exaustivo na elaboração da estratégia de busca.

Por fim, a adoção do *Linked Data* pelas bibliotecas seria vantajosa ao permitir serendipidade a partir dos resultados de busca, ou seja, a descoberta acidental de recursos. (LAURENCE, 2013; ALEMU et al., 2012; CARRAGÁIN et al., 2015; GONZALES, 2014; CATARINO; SOUZA, 2012).

Nesse sentido, entende-se que a adoção do *Linked Data* irá contribuir com as bibliotecas, ao permitir o acesso a dados de fontes externas, que podem complementar a sua descrição dos recursos, criando assim, representações que atendam de maneira mais eficiente às necessidades dos usuários.

Assim, as bibliotecas poderiam contribuir com o *Linked Data* ao disponibilizarem seus dados às demais fontes de recursos informacionais da *Web*, pois a ampla experiência das bibliotecas no estabelecimento de metadados estruturados e padronizados torna-os confiáveis e de interesse para o desenvolvimento do ambiente *Web*.

4 CONSIDERAÇÕES FINAIS

Visando responder à questão de pesquisa “como ocorre a cooperação ente as bibliotecas e demais fontes de informação disponíveis no contexto do *Linked Data*?”, observou-se que essa cooperação pode ser dividida em duas vertentes principais, a de publicadora de dados, onde a maior preocupação é fornecer dados de qualidade; e a de consumidora de dados, onde o desafio é realizar a seleção das fontes a serem utilizadas. A explicitação desse papéis torna-se importante, pois ambos podem ocorrer de maneira independente, exigem tomadas de decisão distintas e implicam em diferentes vantagens.

Os dados das bibliotecas são tradicionalmente estruturados e pautados em instrumentos de padronização que garantem a sua qualidade, entretanto sua utilização é restrita ao domínio bibliográfico. Entende-se que a publicação dos dados bibliográficos como *Linked Data* é vantajosa, pois permitirá seu aproveitamento por instituições externas ao domínio bibliográfico. Enquanto publicadoras de dados, as bibliotecas precisam se preocupar em atender as 35 recomendações apresentadas pelo W3C. A adoção dessas recomendações permite que os consumidores de dados possam identificar suas fontes, selecionar dados de interesse, checar a qualidade, a veracidade e sua consistência e ainda, saber quais as restrições e as permissões de uso dos dados.

A maioria das vantagens identificadas é relacionada às bibliotecas enquanto consumidoras de dados, como a obtenção de dados mais granulares, relação com fontes adicionais de informação do recurso, rompimento das barreiras linguísticas, diminuição do retrabalho e descoberta acidental de recursos. Todavia, isso não se reflete nas recomendações encontradas durante a revisão documental, que são mais direcionadas aos publicadores de dados.

Embora o conjunto de recomendações do W3C (2017) possa auxiliar em ambos os papéis, ao fornecer informações que permitem melhorar a relação entre publicadores e consumidores de dados, ampliando a confiabilidade dos dados disponibilizados de acordo com o *Linked Data*, entende-se que elas não são suficientes para auxiliar no papel dos consumidores de dados, que passariam a depender exclusivamente dos metadados de qualidade fornecidos pelas instituições publicadoras. Nesse sentido, recomenda-se que a comunidade envolvida com o domínio bibliográfico estabeleça instrumentos e critérios que possam embasar sua atuação enquanto consumidora de dados.

REFERÊNCIAS

- AGENJO, X.; HERNÁNDEZ, F.; VIEDMA, A. Data aggregation and dissemination of authority records through linked open data in a european context. **Cataloging & Classification Quarterly**, [s.l.], v. 50, n. 8, p. 803-829, nov. 2012. Disponível em: <http://www.tandfonline.com/doi/abs/10.1080/01639374.2012.711441?journalCode=wccq20>. Acesso em: 15 jul. 2020
- ALEMU, G. et al. Linked data for libraries. **New Library World**, [s.l.], v. 113, n. 11-12, p. 549-570, nov. 2012. Disponível em: <http://www.emeraldinsight.com/doi/full/10.1108/03074801211282920>. Acesso em: 15 jul. 2020.
- ALVES, R. C. V. **Metadados como elementos do processo de catalogação**. 2010. 134 f. Tese (Doutorado em Ciência da Informação) - Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2010.
- ALVES, R. C. V. Metadados editoriais e livreiros: algumas considerações e relações com os padrões de metadados do domínio bibliográfico. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 19, 2018, Londrina. **Anais...** Londrina: Ancib, 2018. p. 5205 - 5219. Disponível em: <http://enancib.marilia.unesp.br/index.php/XIXENANCIB/xixenancib/paper/viewFile/1574/1721>. Acesso em: 15 jul. 2020.
- ARAKAKI, F. A. **Linked data: ligação de dados bibliográficos**. 2016. 144 f. Dissertação (Mestrado em Ciência da Informação) – Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2016.
- BERNERS-LEE, T. **Linked data**, 2006. Disponível em: <https://www.w3.org/DesignIssues/LinkedData.html>. Acesso em: 15 jul. 2020.
- BORIE, J.; VAN BALLEGOOIE, M. From record-bound to boundless: FRBR, linked data, and new possibilities for serials cataloging. **The Serials Librarian**, [s.l.], v. 66, n. 1-4, p. 76-87, maio 2014. Disponível em: <http://www.tandfonline.com/doi/full/10.1080/0361526X.2014.879527.1>. Acesso em: 5 jul. 2020.
- BUSH, V. As we may think. **Atlantic Monthly**, [s.l.], vol.176, n.1, p.101-108, 1945.
- CARRAGÁIN, E. N. Ó. et al. Linked data authority records for Irish place names. **International Journal On Digital Libraries**, [s.l.], v. 15, n. 2-4, p. 74-85, out. 2014. Disponível em: <https://link.springer.com/article/10.1007/s00799-014-0129-8>. Acesso em: 15 jul. 2020.
- CASTRO, F. F.; JESUS, A. F. Adequando dados bibliográficos ao linked data: requisitos necessários, vantagens e desafios. **Informação & Sociedade: Estudos**, Paraíba, v. 28, n. 3, p.149-168, dez. 2018. Disponível em: <https://brapci.inf.br/index.php/res/v/109113>. Acesso em: 15 jul. 2020.
- CATARINO; M. E.; SOUZA, T. B. A representação descritiva no contexto da web semântica. **TransInformação**, Campinas, v. 2, n. 24, p. 77-90, maio-ago., 2012. Disponível em: https://www.scielo.br/scielo.php?pid=S0103-37862012000200001&script=sci_abstract&lng=pt. Acesso em: 15 jul. 2020.

DECLARAÇÃO dos Princípios Internacionais de Catalogação. 2016. p. 1-21, Tradução: Marcelo Votto Texeira. Disponível em: https://www.ifla.org/files/assets/cataloguing/icp/icp_2016-pt.pdf. Acesso em: 15 jul. 2020.

DIAS, C. A. Hipertexto: resumo histórico e efeitos sociais. **Ci. Inf.** Brasília, v. 28, n. 3, p. 269-277, dez 1999. Disponível em: https://www.scielo.br/scielo.php?pid=S0100-19651999000300004&script=sci_abstract&lng=pt. Acesso em: 15 jul. 2020.

GODBY, C. J.; SMITH-YOSHIMURA, K. From records to things: managing the transition from legacy library metadata to linked data. **Bulletin Of The Association For Information Science And Technology**, [s.l.], v. 43, n. 2, p.18-23, dez 2016. Disponível em: <https://doi.org/10.1002/bul2.2017.1720430209>. Acesso em: 15 jul. 2020.

GONZALES, B. M. Linking libraries to the web: linked data and the future of the bibliographic record. **Information Technology And Libraries**, [s.l.], v. 33, n. 4, p.10-22, dez. 2014. Disponível em: <https://ejournals.bc.edu/ojs/index.php/ital/article/view/5631>. Acesso em: 15 jul. 2020.

HAN, M. K. Establishing sustainable and scalable workflows for cataloging and metadata services. **Library Management**, [s.l.], v. 37, n. 6-7, p. 308-316, ago. 2016. Disponível em: <http://www.emeraldinsight.com/doi/full/10.1108/LM-04-2016-0031>. Acesso em: 15 jul. 2020.

HEATH, T. T.; BIZER, C. **Linked data**: evolving the web into a global data space. Berlim: Morgan & Claypool, 2011. 136 p. Disponível em: <http://linkeddatabook.com/editions/1.0/#htoc8>. Acesso em: 15 jul. 2020.

LAURENCE, C. M. Linked data and the Library of Congress. **Library Philosophy and Practice**, Nebraska, p. 2-24, set. 2013. Disponível em: http://digitalcommons.unl.edu/libphilprac/?utm_source=digitalcommons.unl.edu/libphilprac/1114&utm_medium=PDF&utm_campaign=PDFCoverPages. Acesso em: 15 jul. 2020.

QUEIROZ, N. G.; ARAUJO, S. A. Catálogos on-line: um breve estudo dos catálogos on-line de acesso público (OPAC'S). **Múltiplos olhares em ciência da informação**, [s.l.], v. 3, n. 2, p.1-10, jun. 2013.

RAMALHO, R. A. S. **Web semântica**: aspectos interdisciplinares da gestão de recursos informacionais no âmbito da ciência da informação. 2006. 120 f. Dissertação (Mestrado em Ciência da Informação) - Faculdade de Filosofia e Ciências, Universidade Estadual Paulista "Júlio de Mesquita Filho", Marília, 2006.

RAMALHO, R. A. S.; MARTINS, P. G. M.; SOUSA, J. L. Evolução das linguagens de marcação: um breve histórico à luz da área de ciência da informação. **Informação@Profissões**, v. 6, n. 2, p. 20-34, 2017. Disponível em: <http://www.uel.br/revistas/uel/index.php/infoprof/article/view/33267>. Acesso em: 15 jul. 2020.

RAMALHO, R. A. S.; SOUSA, J. L. Diretrizes para avaliação de sistemas de organização do conhecimento representados em skos. **Informação & Informação**, v. 24, n. 2, p. 126-138, 2019. Disponível em: <http://www.uel.br/revistas/uel/index.php/informacao/article/view/37986>. Acesso em: 15 jul. 2020.

RAUTENBERG, S. et al. **Guia prático para a publicação de dados abertos conectados na web**. Curitiba: Appris, 2018. 280 p.

SANTARÉM SEGUNDO, J. E.; SERRA, L. G. O catálogo da biblioteca e o linked data. **em Questão**, [s.l.], v. 23, n. 2, p.167-185, 26 abr. 2017. Disponível em: <https://seer.ufrgs.br/index.php/EmQuestao/article/view/67162>. Acesso em: 15 jul. 2020.

SARACEVIC, T. Information science. **Encyclopedia of library and information sciences**, [s.l.], v. 1, n. 1, p. 2570-2585, 7 dez. 2009. Disponível em: <https://tefkos.comminfo.rutgers.edu/SaracevicInformationScienceELIS2009.pdf>. Acesso em: 15 jul. 2020.

SMIT, J. A informação na ciência da informação. **InCID: Revista de Ciência da Informação e Documentação**, v. 3, n. 2, p. 84-101, 13 dez. 2012. Disponível em: <https://www.revistas.usp.br/incid/article/view/48655>. Acesso em: 15 jul. 2020.

WENZ, R. Towards a web of data: applications and experiences. **Jilis.it**, Csia, v. 4, n. 1, p. 404-413, jan. 2013. Disponível em: <https://www.jlis.it/article/view/5509/7919>. Acesso em: 15 jul. 2020.

W3C. **Data on the web best practices**. 2017. Disponível em: <https://www.w3.org/wiki/LinkedData>. Acesso em: 15 jul. 2020.

W3C. **Data on the web best practices: Dataset Usage Vocabulary**. 2016. Disponível em: <https://www.w3.org/TR/vocab-duv/>. Acesso em: 24 jul. 2020

W3C. **Help and FAQ**. 2019. Disponível em: <https://www.w3.org/Help/#Webinternet>. Acesso em: 15 jul. 2020.

W3C. **Links in HTML documents**. 2018. Disponível em: <https://www.w3.org/TR/html401/cover.html#minitoc>. Acesso em: 15 jul. 2020.

W3C. **Primer RDF**. 2004. Disponível em: <https://www.w3.org/TR/rdf-primer/#intro>. Acesso em: 15 jul. 2020.

W3C. **URIs, URLs e URNs: Clarifications and Recommendations 1.0**. 2011. Disponível em: <https://www.w3.org/TR/uri-clarification/>. Acesso em: 15 jul. 2020.

W3C INCUBATOR GROUP. **Library Linked Data incubator group: final report**. 2011. [S.L.]. 19 p. Disponível em: <https://www.w3.org/2005/Incubator/lld/>. Acesso em: 15 jul. 2020.

XU, A. et al. Initial BIBFRAME 2.0 modeling for the library information spotlight “Opera Planet”. **Journal Of Library Metadata**, [s.l.], v. 16, n. 3-4, p.202-227, out. 2016. Disponível em: <http://www.tandfonline.com/doi/full/10.1080/19386389.2016.1258910>. Acesso em: 15 jul. 2020.

Notas

AGRADECIMENTOS

Agradecemos a Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) que possibilitou e financiou a realização dessa pesquisa. (processo n° 2018/26545-7 -DIRETRIZES PARA A ADEQUAÇÃO DE DADOS BIBLIOGRÁFICOS AO LINKED (OPEN) DATA.

CONTRIBUIÇÃO DE AUTORIA

Concepção e elaboração do manuscrito: A. F. de Jesus, F. F. de Castro, R. A. S. Ramalho

Coleta de dados: A. F. de Jesus

Análise de dados: A. F. de Jesus, F. F. de Castro, R. A. S. Ramalho

Discussão dos resultados: A. F. de Jesus, F. F. de Castro, R. A. S. Ramalho

Revisão e aprovação: A. F. de Jesus, F. F. de Castro, R. A. S. Ramalho

CONJUNTO DE DADOS DE PESQUISA

Todo o conjunto de dados que dá suporte aos resultados deste estudo foi publicado no próprio artigo.

FINANCIAMENTO

Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP - processo n° 2018/26545-7)

CONSENTIMENTO DE USO DE IMAGEM:

Não se aplica.

APROVAÇÃO DE COMITÊ DE ÉTICA EM PESQUISA

Não se aplica.

CONFLITO DE INTERESSES

Não se aplica.

LICENÇA DE USO

Os autores cedem à **Encontros Bibli** os direitos exclusivos de primeira publicação, com o 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. Os **autores** têm autorização para assumir contratos adicionais separadamente, para distribuição não exclusiva da versão do trabalho publicada neste periódico (ex.: publicar em repositório institucional, em site pessoal, publicar uma tradução, ou como capítulo de livro), com reconhecimento de autoria e publicação inicial neste periódico.

PUBLISHER

Universidade Federal de Santa Catarina. Programa de Pós-graduação em Ciência da Informação. Publicação no [Portal de Periódicos UFSC](#). As ideias expressadas neste artigo são de responsabilidade de seus autores, não representando, necessariamente, a opinião dos editores ou da universidade.

EDITORES

Enrique Muriel-Torrado, Edgar Bisset Alvarez, Camila Barros.

HISTÓRICO

Recebido em: 23/07/2020 – Aprovado em: 23/11/2020 –Publicado em: 04/01/2021