

XVIII ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO – ENANCIB 2017

GT-8 – Informação e Tecnologia

O ORCID COMO APLICAÇÃO DE LINKED DATA NO CATÁLOGO DE BIBLIOTECAS

Liliana Giusti Serra – (Universidade Estadual Paulista – UNESP)

Luciana Candida da Silva – (Universidade Estadual Paulista – UNESP)

José Eduardo Santarém Segundo – (Universidade Estadual Paulista – UNESP)

THE ORCID AS A LINKED DATA APPLICATION ON THE LIBRARY CATALOG

Modalidade da Apresentação: Comunicação Oral

Resumo: As tecnologias da Web Semântica, por meio do *Linked Data*, proporcionam o reaproveitamento de dados existentes na Web e enriquece as informações do catálogo das bibliotecas, ampliando a oferta de informações disponibilizadas aos usuários. O objetivo deste estudo é analisar a possibilidade do *dataset* ORCID ser utilizado nos catálogos das bibliotecas para contribuir com a descrição de dados de autoridades, além de controlar ambiguidades de nomes de autores pessoais presentes no ambiente acadêmico. Trata-se de uma pesquisa de natureza qualitativa, do tipo descritiva-exploratória, que descreve a evolução do catálogo e seus princípios e, posteriormente, os elementos da Web Semântica, com foco na aplicação do *Linked Data*. O ORCID foi detalhado em seus objetivos e evolução, com ênfase em seus aspectos semânticos. Em seguida, foram realizados levantamentos e paridades dos elementos semânticos presentes em registros do ORCID com os campos utilizados para representação de autoridades no formato MARC 21. Como resultado, obteve-se que os dados coletados no ORCID podem ser reutilizados no controle de autoridades evitando ambiguidade de nomes, facilitando o cadastro e manutenção de nomes de pessoas, além de permitir aos usuários a descoberta de outras produções científicas realizadas por esta autoria que não fazem parte do acervo da biblioteca, mas que podem ser localizadas em outros sítios, por meio de URIs. Esta iniciativa mostra que é possível aplicar o *Linked Data* para o reaproveitamento de dados de autoridades, por meio de vínculos, agregando informações aos catálogos das bibliotecas.

Palavras-Chave: Catálogo de Bibliotecas; *Linked Data*; Web Semântica; *Dataset* ORCID; Controle de Autoridades.

Abstract: Linked data, as a Semantic Web technology, provides the reuse of existing data on the Web and enrich the information in the library catalog, expanding the range of information available to patrons. The purpose of this study is to analyze the possibility of the ORCID dataset being used in library catalogs to contribute to the description of authority data, as well as to control ambiguities of personal author names existing in the academic environment. This is a qualitative, descriptive-exploratory research that describes the evolution of the catalog and its principles, and furthermore, the elements of the Semantic Web, focusing on the application of Linked Data. The ORCID was detailed in its

objectives and evolution, with emphasis on its semantic aspects. Subsequently, analysis and parity of the semantic elements present in ORCID records, identifying the fields used to represent authorities in the MARC 21 format were performed. As a result, it was observed that the data collected in the ORCID can be reused in the control of authorities, facilitating the registration and maintenance of names of authors and controlling ambiguities, as well as allowing users to discover other scientific productions made by those authors, which are not present in the collection of the library, but which can be located in other places, through URIs. This initiative shows that it is possible to apply Linked Data for the reuse of authority data, through links, aggregating information to the libraries' catalog.

Keywords: Library Catalog; Linked Data; Semantic Web; ORCID Dataset; Control of Authorities.

1 INTRODUÇÃO

Os catálogos das bibliotecas proporcionam informações detalhadas e estruturadas dos recursos informacionais presentes nos acervos. Assim, o usuário pode consultar se determinada obra pertence à coleção, se está disponível e quais as condições para sua utilização. A representação é realizada por meio de aplicação de regras que norteiam como um recurso deve ser descrito de forma a ser localizado pelo usuário de variadas formas, como pelo autor, título, assuntos etc.

Enquanto os catálogos estavam restritos ao espaço da biblioteca, os usuários somente conseguiam fazer consultas localmente, deslocando-se até o espaço para descobrir se a biblioteca possuía um exemplar de uma obra em particular e, em caso afirmativo, se a mesma estaria disponível para leitura. Ao publicar os catálogos na Web, a biblioteca passa a permitir aos usuários a consulta de forma remota, com o deslocamento do usuário ocorrendo somente se a obra existe e se pode ser consultada. Entretanto, ao dispor do catálogo na Web, outras possibilidades de uso transcendem a forma tradicional, tendo alcance maior em relação às fichas impressas, podendo ser consultado de qualquer local, divulgando as obras presentes em um acervo. Mas o uso e descoberta dos catálogos continua restrito ao acervo da instituição, agora ao seu sítio, sem permitir a frequência de dados complementares que podem incluir informações de forma dinâmica, enriquecendo o catálogo, auxiliando na contextualização de pesquisas e proporcionando acesso a outros recursos, além dos presentes no acervo.

Não basta, portanto, publicar o catálogo na Web e deixá-lo legível somente por pessoas. É necessário que o catálogo seja legível também por máquinas, onde os dados são identificados, reconhecidos e reutilizados em diversas aplicações. As tecnologias da Web Semântica permitem, por meio do *Linked Data* (LD), que informações dispersas na Web sejam interligadas, favorecendo novas descobertas e reutilização de informações, diminuindo ou eliminando a criação de dados em detrimento do reaproveitamento dos mesmos.

Dentre as aplicações de LD, neste estudo será analisado o *Open Researcher and Contributor ID* (ORCID) como um *dataset* que pode ser utilizado nos catálogos das bibliotecas para contribuir com a descrição e enriquecimento de dados de autoridades. Serão identificados os dados presentes no ORCID

que podem ser reutilizados no cadastro da autoridade, facilitando as atividades da biblioteca e contribuindo com a oferta de informações com maior detalhamento aos usuários.

2 PROCEDIMENTOS METODOLÓGICOS

O artigo é fruto de pesquisa de natureza qualitativa, do tipo descritiva-exploratória, e busca analisar as possibilidades de vinculação do *dataset* ORCID ao catálogo da biblioteca para reaproveitamento e enriquecimento dos metadados do cadastro de autoridades. Para tanto, primeiramente é descrita a evolução do catálogo e seus princípios e, posteriormente, são identificados elementos da Web Semântica, com foco na aplicação do LD. O ORCID, por tratar-se de um *dataset* estruturado, foi a aplicação de LD selecionada para este estudo. Assim, o ORCID foi detalhado em seus objetivos e evolução, com ênfase em seus aspectos semânticos. Após identificação dos elementos semânticos presentes no ORCID, foram realizados levantamentos e paridades destes com os campos utilizados para representação de autoridades no formato MARC, padrão utilizado pelas bibliotecas.

A literatura consultada foi selecionada nas bases de dados *Scopus*¹, *Brapci*² e Portal de Periódicos Capes³. Também foram pesquisados livros, artigos de revistas e sítios da Web. O recorte temporal utilizado possui ênfase no período de 2010 a 2017 no que concerne à Web Semântica, LD e ORCID. Esta literatura é formada, essencialmente, por produção estrangeira. Este fator reforça a importância em pesquisas sobre o tema. O referencial teórico sobre catálogos corresponde ao período de 1969 a 2017, abrangendo obras consideradas clássicas até as discussões mais recentes.

3 O CATÁLOGO

O catálogo percorreu uma longa trajetória que se permeia pela formulação de princípios, elaboração de regras para descrição bibliográfica, escolha de pontos de acesso, definição de formatos para intercâmbio de dados e estruturação de modelos conceituais para dados bibliográficos e de autoridade. As funções e a estrutura do catálogo atual tiveram sua base construída a partir dos princípios delineados por teóricos da catalogação desde a metade do século XIX.

Em 1841 Anthony Panizzi aprovou 91 regras de catalogação para compilação do catálogo do Museu Britânico. Nessas regras incluía a concepção de que a construção de um catálogo envolvia a compreensão da obra alinhada aos interesses dos usuários. A visão da catalogação com base na obra é retomada amplamente na década de 1990 com o *Functional Requirements of Bibliographic Records* (FRBR), modelo conceitual do tipo entidade-relacionamento, e o *Functional Requirements for Authority* (FRAD), para análise conceitual dos dados de autoridade.

¹ Scopus: <https://www.scopus.com/>

² Brapci: <http://www.brapci.ufpr.br/brapci/>

³ Portal de Periódicos Capes: <http://www.periodicos.capes.gov.br/>

**XVIII ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO – ENANCIB 2017
23 a 27 de outubro de 2017 – Marília – SP**

Os objetivos do catálogo foram determinados por Charles Ammi Cutter, em 1876, com a publicação da obra *Rules for a dictionary catalog*. Nesta obra o catálogo é identificado como um instrumento que permite encontrar um livro do qual se conheça o autor, o título e/ou o assunto, além de mostrar o que existe em uma coleção (BARBOSA, 1978). Para Lubetzky (1969), Cutter não qualificou nos objetivos do catálogo a distinção entre livro e obra. Segundo Lubetzky (1969), o livro é o registro material e a obra é o produto intelectual incorporado no livro. De acordo com Mey e Silveira (2009), os objetivos do catálogo definidos por Cutter e revisados por Lubetzky foram os mesmos especificados na Conferência de Paris, em 1961, e são aceitos até os dias de hoje. A Conferência Internacional sobre Princípios de Catalogação, realizada em 1961, também chamada de Conferência de Paris, foi um importante passo para a consolidação dos princípios do catálogo, que foram confirmados e atualizados pela Declaração de Princípios, realizada em 2003, pela *International Federation of Library Associations and Institutions* (IFLA).

Para Santos e Pereira (2014), Jewett ao propor em 1852 a reprodução das fichas catalográficas do catálogo da biblioteca da *Smithsonian Institution* para serem trocadas com outras bibliotecas, preconizava a catalogação cooperativa. Entretanto, as autoras ponderam que a catalogação cooperativa veio a se efetivar com a adoção do *Machine Readable Cataloging* (MARC), formato de intercâmbio de dados, desenvolvido pela *Library of Congress*, na década de 1960.

O catálogo enquanto meio de comunicação entre o usuário e a informação é o instrumento mais antigo das bibliotecas, o qual pode ser compreendido como:

[...] um meio de comunicação que veicula mensagens sobre os registros do conhecimento, de um ou vários acervos, reais ou ciberespaciais, apresentando-as com sintaxe e semânticas próprias e reunindo os registros do conhecimento por semelhanças, para os usuários desses acervos. O catálogo explicita, por meio de mensagens, os atributos das entidades e os relacionamentos entre elas. (MEY; SILVEIRA, 2009, p.12).

A Declaração de Princípios de 2003 foi baseada nos modelos conceituais FRBR e FRAD e nas tradições catalográficas para atender aos objetivos do catálogo. De acordo com a IFLA (2009) o primeiro princípio desta Declaração é servir a conveniência dos usuários, seguido dos demais princípios como a consistência e a padronização, onde as descrições e a construção de pontos de acesso devem ser uniformizadas para aumentar a capacidade de compartilhar dados bibliográficos e de autoridade.

Um formato de metadados muito usado em sistemas de bibliotecas é o MARC 21, utilizando a estrutura de campos, subcampos e indicadores para descrição. Este formato permite estruturar dados de forma detalhada e com qualidade para descrição dos recursos. O formato MARC, entretanto, mesmo passando por adaptações em sua estrutura, tem-se mostrado inapropriado para atender as potencialidades dos novos modelos e padrões aplicados ao domínio bibliográfico. Contudo, acredita-

**XVIII ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO – ENANCIB 2017
23 a 27 de outubro de 2017 – Marília – SP**

se que as bibliotecas continuem usando este formato nos próximos anos, pelo menos enquanto seu substituto, *Bibliographic Framework* (BIBFRAME), não estiver totalmente disponível para implantação.

Os padrões descritivos utilizados em bibliotecas são o *Anglo American Cataloguing Rules 2nd Edition* (AACR2), o *International Standard Bibliographic Description* (ISBD) e o *Resource Description and Access* (RDA). Estes instrumentos proporcionam regras para realização de descrições que atendem aos usuários e possibilitam a recuperação dos dados. Segundo Coyle (2010), os dados dos catálogos foram planejados para serem consumidos por pessoas e não por máquinas. Dispor dos dados dos catálogos na Web para que também possam ser utilizados por máquinas e oferecidos como serviços aos usuários, aumenta as chances de descoberta e reutilização de informações. Ao reaproveitar dados existentes na Web para enriquecer o catálogo por meio de relacionamentos com outras fontes externas, permite-se a ampliação de utilização em contextos variados, afinal “A Web fornece uma plataforma para ligar recursos de informação independente de sua proveniência” (COYLE, 2010, p. 7, tradução nossa).

A adoção de padrões por si só não garante a interconexão dos dados de autoridades e bibliográficos de um recurso para o outro apresentando as informações com sintaxe e semântica, como proposto nos princípios e funções do catálogo. Para tanto, a implementação do LD como a aplicação de conceitos e tecnologias da Web Semântica apresenta-se como a possibilidade para atender a proposta do catálogo atual.

4 LINKED DATA

As informações contidas nos catálogos das bibliotecas são legíveis por humanos, mas não otimizadas para a compreensão por computadores. No contexto atual, o texto é colocado na Web sem marcação que aporte significado estrutural e semântico, o que impossibilita a ligação automática de um recurso a outro (HASTINGS, 2015).

Para Berners-Lee (2006), a Web Semântica consiste em fazer ligações, de modo que uma pessoa ou máquina possa explorar a Web de dados. Segundo Bizer, Heath e Berners-Lee (2009), tradicionalmente, os dados publicados na Web foram disponibilizados com marcações *HyperText Markup Language* (HTML), sacrificando grande parte de sua estrutura semântica, pois o HTML não é suficientemente expressivo para permitir que entidades individuais descritivas em um documento sejam conectadas por *links* a entidades relacionadas. A Web Semântica é a visão do *World Wide Web Consortium* (W3C⁴) sobre a Web de dados conectados, ou LD, concebido para padronizar a conexão entre dados. Para Bizer, Heath e Berners-Lee (2009, p. 2, tradução nossa), “*Linked Data* é simplesmente como usar a Web para criar ligações digitadas entre dados de diferentes fontes”.

⁴ W3C: <https://www.w3.org/>

XVIII ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO – ENANCIB 2017
23 a 27 de outubro de 2017 – Marília – SP

O LD segue um conjunto de princípios projetados em 2006 por Tim Berners-Lee para promover a vinculação entre os dados na Web, facilitando a ligação entre recursos e sua reutilização. Estes princípios são:

1. Usar *Uniform Resource Identifier* (URIs) para nomear itens;
2. Usar URIs HTTP para que pessoas possam procurar esses nomes;
3. Ao consultar uma URI, fornecer informações em formatos padronizados úteis [*Resource Description Framework* (RDF), SPARQL etc.];
4. Incluir sentenças com *links* para outras URIs, para permitir que itens relacionados possam ser descobertos (BERNERS-LEE, 2006).

Assim, o LD pode ser definido como “um conjunto de boas práticas para publicação e conexão de dados estruturados na Web, usando padrões internacionais recomendados pelo W3C” (ISOTANI; BITTENCOURT, 2015, p. 14).

Um dos padrões comumente recomendado pela W3C para ser usado em projetos de LD é o RDF, modelo de dados que codifica o significado de etiquetas por meio de triplas, formadas por recurso, propriedade e valor. O recurso é um objeto da Web identificado por uma URI. A propriedade é uma característica, um atributo ou uma relação usada para descrever um recurso. O valor é o dado atribuído ao recurso por meio da propriedade. No contexto do catálogo, uma tripla pode representar a relação de autoria entre pessoa e obra. Assim, por exemplo, Machado de Assis (recurso) é autor (propriedade) da obra Quincas Borba (valor). As triplas são vinculadas entre si através do uso de URIs, visto que identificam recursos e expressam as relações entre eles. Segundo Ryan et al. (2015), usar uma URI em vez de uma sequência de texto para identificar um recurso remove qualquer ambiguidade que exista entre termos que tenham a mesma descrição sintática.

A Web Semântica utiliza diversas tecnologias. Os formatos mais utilizados são o *extensible Markup Language* (XML), Turtle, N-Triples e *JavaScript Object Notation* (JSON). As declarações dos recursos podem ser estruturadas pelo *RDF Schema* (RDFS) ou pela *Web Ontology Language* (OWL). Ontologias legíveis por máquinas podem ser aplicadas, como é o caso da *Friend of a Friend* (FOAF) e o *Simple Knowledge Organization System* (SKOS). O FOAF é utilizado para ligar pessoas e informações. Já o SKOS desenvolve especificações e padrões para utilização de sistemas de organização do conhecimento, como tesouros, esquemas de classificação, taxonomias e cabeçalhos de assuntos no âmbito da Web Semântica. O SPARQL é um conjunto de especificações que fornecem linguagens e protocolos para recuperar e manipular dados armazenados em RDF (W3C, 2013). Ao implantar tais tecnologias, os dados podem ser publicados na Web em padrões que são entendidos semanticamente pelos computadores. Ao criar *links* entre diversos conjuntos de dados, o LD quebra as barreiras entre eles, permitindo descoberta, compartilhamento e reutilização de dados (RYAN et al., 2017).

XVIII ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO – ENANCIB 2017 23 a 27 de outubro de 2017 – Marília – SP

A interconexão de dados na Web, entretanto, não significa que eles precisam ser abertos. Segundo a *Open Knowledge International*⁵ (2004), o *Linked Open Data* (LOD) refere-se aos dados que podem ser livremente utilizados, reutilizados e redistribuídos por qualquer pessoa, contanto que seja feita atribuição da fonte original e compartilhamento pelas mesmas licenças em que as informações foram apresentadas. A interligação de dados pode ocorrer também com dados fechados.

Pensar no LD no contexto da biblioteca pode ser uma oportunidade para enriquecer os dados do catálogo com outras fontes, e assim ampliar a oferta de informações disponibilizadas aos usuários. Para Coyle (2010), a biblioteca encontra-se no momento oportuno para vislumbrar a interconexão dos dados do catálogo, pois possui metadados estruturados, descritos com base em normas e padrões que permitem a sua codificação com alto nível de detalhamento, o que favorece a transformação do modelo de registro para um modelo de dados.

Os padrões emergentes do domínio bibliográfico estabelecem base para implantação do LD em bibliotecas. Os modelos conceituais FRBR e FRAD se apresentam em grupos de entidades e relacionamentos e serviram de base para o desenvolvimento do código RDA. O BIBFRAME é uma estrutura em desenvolvimento, iniciado pela *Library of Congress* em 2011, para descrição bibliográfica baseada em LD. De acordo com a *Library of Congress* (2012), o BIBFRAME promove a prática de modelagem do RDF que identifica como recursos da Web todas as entidades, atributos e relações. Isso permite que mais anotações, como mapeamentos para outros vocabulários ou extensões da comunidade local, sejam criadas. Os dados bibliográficos enriquecidos com o LD podem proporcionar outras oportunidades de navegação e oferta de pontos de acesso aos usuários. Com os princípios do LD, os metadados das bibliotecas também podem ser encontrados pelos motores de busca, transformando catálogos estáticos em repositórios dinâmicos, formados pelas informações representadas pelas bibliotecas e enriquecidas por outras fontes da Web.

5 ORCID

O ORCID é uma instituição sem fins lucrativos, internacional e interdisciplinar, fundada em 2010, que visa criar e manter registro de pesquisadores, autores acadêmicos e demais contribuintes, provendo um identificador único a cada um, de forma a conectar atividades, produções e resultados de pesquisa, em âmbito mundial. O projeto do ORCID foi iniciado em 2009, liderado pela *Thompson Reuters* e o grupo editorial *Nature*. Estes parceiros vislumbravam criar uma organização sem interesses financeiros para identificar, de forma aberta e gratuita, os autores de produções acadêmicas em todas as disciplinas, em nível nacional e internacional, e que permitisse interação com outros sistemas de identificação de autores. Em 2012 o serviço não proprietário foi lançado, com abordagem cooperativa, de forma gratuita e licença de código aberto (WALTON, 2014).

⁵ *Open Knowledge International*: <https://okfn.org/>

XVIII ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO – ENANCIB 2017
23 a 27 de outubro de 2017 – Marília – SP

Seus objetivos centram-se em dirimir problemas de ambiguidades de nomes em comunicações científicas, gerenciar publicações, estabelecer vínculo entre atividades de pesquisa e resultados, com vistas a facilitar a colaboração em rede (CASSOTTA et al., 2017). As informações dos autores presentes no ORCID reúnem dados biográficos e bibliográficos, estes últimos podendo conter outros identificadores como o DOI (*Digital Object Identifiers*), ou textos abertos de publicações. Os perfis dos autores podem ser utilizados para identificar conteúdos, aumentando as chances de descoberta de produções acadêmicas, mas também para acompanhar métricas (WILSON; FENNER, 2012). Seus princípios norteiam-se pela criação de registro permanente, com atribuição confiável de produção, superando barreiras disciplinares, geográficas, nacionais e internacionais. Os dados presentes no ORCID são disponibilizados essencialmente de forma gratuita, mas também dispõem de serviços pagos.

Sua aplicação se dá no meio acadêmico, permitindo que autores sejam identificados de forma única por meio de um código, independente da forma como os nomes foram grafados em artigos, livros, relatórios de pesquisas, teses e demais produções acadêmicas. Ao proporcionar a identificação de autores por meio de um código, possibilita que variações de nomes sejam superadas, sejam elas decorrentes de homônimos, alterações devido a casamentos ou separações e divórcios, diferenças culturais, abreviações de prenomes ou outras formas que resultam em dificuldades para identificar, recuperar ou reunir dados e produções de pesquisadores.

A comunidade ORCID é formada por pesquisadores, editores, universidades, governos, centros de pesquisa, laboratórios, agências de fomento, repositórios de dados, sociedades de profissionais etc. A criação de um código ORCID é gratuita para pesquisadores, enquanto organizações pagam taxas com custos variados, de acordo com seu escopo e tamanho.

O código ORCID é uma URI formada pelo prefixo <http://orcid.org/> seguido de conjunto de dezesseis caracteres, essencialmente numéricos, separados por hífen a cada grupo de quatro dígitos, como, por exemplo, <http://orcid.org/xxxx-xxxx-xxxx-xxxx>, sendo que o último dígito pode ser uma letra, funcionando como um dígito de verificação, característica existente em outros identificadores únicos como, por exemplo, o *International Standard Book Number* (ISBN). Explicações sobre a formação do código, sua segmentação, tamanho, verificação e acessibilidade são relatadas por Bilder (2012). O ORCID utiliza um formato compatível com o *International Standard Name Identifier* (ISNI) e adota o padrão da norma ISO 27729.

O código ORCID pode ser empregado no processo de submissão de manuscritos (textos para congressos, revistas etc.), permitindo agrupar dados em ponto único, mesmo que outros identificadores sejam utilizados pelos autores em âmbito regional. Isto contribui com a identificação da autoridade e demais colaboradores na origem do processo de publicação, favorecendo acompanhamento e controle de textos, similaridades de pesquisas, conflitos de interesse, atribuições de revisões etc.

XVIII ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO – ENANCIB 2017 23 a 27 de outubro de 2017 – Marília – SP

A interoperabilidade é parte fundamental do ORCID. Assim, a possibilidade de ser integrado com outros sistemas, como por exemplo, nos processos de submissão de trabalhos acadêmicos e demais atividades de pesquisa como publicações, prêmios, patentes e *datasets*, proporciona facilidades aos pesquisadores. Ao se identificar pelo seu código ORCID, o pesquisador não precisa informar os dados de nome, afiliação ou *e-mail*, visto que são fornecidos de forma automática. Também é possível a criação de relação entre autores e editores onde, ao ter um texto aceito, os metadados do artigo serão atualizados na publicação e na produção do autor. Os pesquisadores também podem gerenciar seus cadastros, definindo aspectos de privacidade de seus dados, como quais informações podem ser públicas ou de acesso restrito, ou o que pode ser compartilhado com outros parceiros, identificados e definidos pelo autor (HAAK et al, 2012). O código identificador é obrigatoriamente público, porém as demais informações, caso tenham sido definidas como privadas, não são compartilhadas. Os dados de afiliação, país, instituição, área de estudo etc. podem ser alterados no cadastro, de acordo com o andamento da carreira do pesquisador (ORCID, 2017b), sem interferir em sua identificação ou código.

Segundo Haak et al. (2012), ao utilizar o código ORCID para identificar autores, além das questões de ambiguidade de nomes, é facilitada a criação e atualização de lista de publicações. Isto é particularmente sensível em repositórios digitais, onde a inclusão de novas produções pode ser realizada pelo próprio autor, pela biblioteca, ou ainda pelo editor. Ao identificar o autor pelo seu código e não pelo seu nome, dados de produções e financiamentos podem ser recuperados e reunidos com menor esforço, permitindo agrupar e rastrear as informações de pesquisas em andamento. Ao centralizar os dados em um único ponto e existindo a interoperabilidade com outros sistemas, é dispensada a necessidade de realização de múltiplos cadastros ao submeter manuscritos, ou ainda em atualizar dados em produções anteriores, visto que a identificação e localização dos autores passa a ser persistente. Outro princípio do ORCID é prover os dados de forma aberta e gratuita por meio de *Application Programming Interface* (APIs) que suportam a comunicação e autenticação com outros sistemas, disponibilizando dados em licença de fonte aberta (ORCID, 2017).

De acordo com Carter e Blandford (2017), apesar de existirem outros sistemas que visam identificar autores acadêmicos como *ResearcherID*, *Scopus Author ID* e *Google Scholar*, o ORCID não visa lucros e não utiliza plataforma proprietária, poupa o tempo do pesquisador ao preencher seus dados em diversos sítios da Web, permite vincular dados dos autores a outros sistemas por meio de APIs e possibilita a reunião da produção, incluindo fontes alternativas, consideradas não acadêmicas, como participação em programas de televisão ou *tweets*.

5.1 Aspectos Semânticos do *Orcid*

De acordo com Bilder (2012), o desenvolvimento do ORCID teve como impute original a estrutura utilizada no *ResearcherID*, plataforma proprietária desenvolvida pela *Thompson Reuters*. Ao utilizar-se

XVIII ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO – ENANCIB 2017
23 a 27 de outubro de 2017 – Marília – SP

desta plataforma, alterações na estrutura dos dados se fizeram necessárias com o intuito de, primeiramente, tornar a ferramenta completamente aberta, alterando características específicas do sistema utilizado no *ResearcherID*. Este esforço acarretou em desambiguação entre o *ResearcherID* e o ORCID, visando construir uma base neutra, não vinculada com serviços proprietários, inclusive para não conflitar com os perfis existentes na base da *Thompson Reuters*. Outro aspecto era a amplitude da ferramenta. A estrutura do *ResearcherID* conta com identificadores com 11 a 13 caracteres o que, a curto prazo, poderia resultar em esgotamento de códigos identificadores disponíveis.

Também decorria do excesso de informações semânticas que podem ser incluídas no *ResearcherID*, o que inviabiliza a adoção de um código único que, espera-se, seja persistente em longo termo. Dados como nomes, *e-mails*, instituições, afiliação, áreas de estudos, financiamento de projetos etc., mudam constantemente e, portanto, não podem estar diretamente vinculados ao código, mas devem ser descritos como atributos do pesquisador. Manter uma estrutura onde atributos do pesquisador fazem parte do código identificador resulta em enorme complexidade para manter uma identificação única, visto que um autor, no decorrer de sua carreira, pode mudar de instituição de pesquisa, disciplinas, interesses de estudo, país de atuação etc. Para tanto, é necessário um identificador que permita o reconhecimento único do pesquisador, porém sem a vinculação de seus atributos a este código. Com o intuito de tornar os identificadores do ORCID semanticamente opacos decidiu-se que o código seria formado somente por números, gerados de forma não sequencial e sem qualquer significado semântico que permita a identificação do autor, instituição afiliada, vínculo com editores ou localização geográfica.

De acordo com o DataHub (2017), o *dataset* do ORCID utiliza o modelo RDF com vocabulários bem conhecidos como *Dublin Core*, *FOAF*, *Schema.org*, *Geonames*, entre outros. Os dados descritivos de autores presentes no ORCID estão agrupadas em categorias como:

(a) Dados biográficos: nome, formas variantes do nome, país, palavras-chave, *e-mails*, *websites*, outros identificadores etc.;

(b) Educação: instituição, data inicial e final (dia, mês e ano), localização (cidade, estado, região, país), título obtido etc.;

(c) Emprego: instituição, data inicial e final, localização, nome da organização e departamento, cargo etc.;

(d) Produções: título, idioma, ISSN, dados do fascículo, identificadores (DOI), tipo de publicação etc.;

(e) Financiamentos: tipo de financiamento, projeto financiado, nome da agência de fomento, valor, URL, tipo de relação, localização da agência de fomento, número e URL da bolsa concedida etc.

XVIII ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO – ENANCIB 2017
23 a 27 de outubro de 2017 – Marília – SP

O *dataset* ORCID dispõem os dados em HTML, RDF/XML, RDF Turtle, XML e JSON. Os dados utilizados neste estudo foram coletados por meio da aplicação *EasyRDF Converter*⁶, no formato *RDF Turtle*. Foi escolhido, aleatoriamente, o pesquisador Rami Albatal (ORCID ID 0000-0002-9269-8578), da Irlanda. Para efeitos de conferência do RDF obtido, o registro utilizado como exemplo também foi extraído por meio de linha de comando (`curl -v -L -H "Accept: text/turtle" https://orcid.org/0000-0002-9269-8578`). As informações coletadas pelo *EasyRDF* e pela linha de comando foram as mesmas e estão dispostas no Quadro 1. As categorias educação (*education*), emprego (*employment, activities*), financiamento (*funding*) e produções (*works*), apesar de interessantes às bibliotecas, serão abordadas em estudos futuros.

Quadro 1: Dados coletados do ORCID em RFD/Turtle

```
@prefix foaf: <http://xmlns.com/foaf/0.1/> .
@prefix ns0: <http://purl.org/pav/> .
@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .
@prefix prov: <http://www.w3.org/ns/prov#> .
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
@prefix ns1: <http://www.geonames.org/ontology#> .

<http://orcid.org/0000-0002-9269-8578#workspace-works> a foaf:Document .
<http://pub.orcid.org/orcid-pub-web/experimental_rdf_v1/0000-0002-9269-8578>
  a foaf:PersonalProfileDocument ;
  ns0:createdBy <http://orcid.org/0000-0002-9269-8578> ;
  ns0:createdOn "2015-01-27T23:00:34.956Z"^^xsd:dateTime ;
  ns0:createdWith <http://orcid.org> ;
  ns0:lastUpdateOn "2016-09-21T01:50:38.178Z"^^xsd:dateTime ;
  prov:generatedAtTime "2016-09-21T01:50:38.178Z"^^xsd:dateTime ;
  prov:wasAttributedTo <http://orcid.org/0000-0002-9269-8578> ;
  foaf:maker <http://orcid.org/0000-0002-9269-8578> ;
  foaf:primaryTopic <http://orcid.org/0000-0002-9269-8578> .

<http://orcid.org/0000-0002-9269-8578#orcid-id>
  a foaf:OnlineAccount ;
  rdfs:label "0000-0002-9269-8578" ;
  foaf:accountName "0000-0002-9269-8578" ;
  foaf:accountServiceHomepage <http://orcid.org> .

<http://sws.geonames.org/2963597/>
  rdfs:label "Ireland", "Éire" ;
  ns1:countryCode "IE" ;
  ns1:name "Ireland", "Éire" .

<http://orcid.org/0000-0002-9269-8578>
  a prov:Person, foaf:Person ;
  rdfs:label "Rami Albatal" ;
  foaf:account <http://orcid.org/0000-0002-9269-8578#orcid-id> ;
  foaf:based_near [
    a ns1:Feature ;
    ns1:countryCode "IE" ;
    ns1:parentCountry <http://sws.geonames.org/2963597/>
  ] ;
  foaf:familyName "Albatal" ;
  foaf:givenName "Rami" ;
  foaf:name "Rami Albatal" ;
  foaf:page <http://ie.linkedin.com/in/ramialbatal>,
<http://scholar.google.co.nz/citations?user=tuG6AlQAAA&hl=en> ;
  foaf:plan ""Data Scientist, Ph.D. graduated in 2010 from Grenoble University - France.
```

⁶ *Easy Converter*: <http://www.easyrdf.org/converter>

**XVIII ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO – ENANCIB 2017
23 a 27 de outubro de 2017 – Marília – SP**

```

Sing April 2012, I am the lead postdoctoral researcher the Lifelogging team at Insight
Centre for Data Analytics - Dublin City University.
My research focuses on three main area: Lifelogging, Automatic Image Classification and
Machine Learning. After having achieved my Ph.D. and worked for two years on research
related to visual content analysis and retrieval, I am currently working on proposing new
generation of Lifelogging technologies that employ advanced sensed data analysis and
machine learning, in the goal of building external human memory that supports the total
recall and mining and improve data-derived decision making and planning.
I am actively involved in organizing international conferences, reviewing high-level
scientific journals, and seeking national, European and International funding opportunities
for research projects."";
foaf:publications <http://orcid.org/0000-0002-9269-8578#workspace-works> .
    
```

Fonte: Dados retirados do ORCID e visualizados no padrão RDF/Turtle pelo Easy Converter - 2017.

Verificou-se que é possível a coleta de informações biográficas (*person, other names*), para reaproveitamento de dados no catálogo de autoridades, porém nem todas as informações visíveis no sítio constavam no RDF, apesar de estarem visíveis por pessoas no sítio do ORCID. Este fato, provavelmente, é decorrente do estágio experimental do projeto, com discussões sendo realizadas a fim de melhorar a oferta de dados disponíveis no *dataset* (DEMÉRANVILLE, 2017; ARCHER, 2015).

5.2 Elementos do *Orcid* e o Formato MARC

Após a coleta dos dados presentes no RDF foram elencados os campos do MARC autoridades e bibliográfico onde estes dados podem ser reaproveitados. A identificação dos campos, indicadores e subcampos do MARC foram realizadas de acordo com os dados presentes no sítio da *Library of Congress*, instituição que mantém e atualiza o formato (LIBRARY OF CONGRESS, 2017). Assim, pretende-se avaliar a otimização de dados básicos e essenciais de autoridades com os elementos dispostos no ORCID. Os elementos selecionados estão dispostos no Quadro 1.

Quadro 1: Representação de elementos do ORCID no MARC

Elemento	ORCID	MARC Autoridade	MARC Bibliográfico
Nome	foaf:familyName "Albatal" ; foaf:givenName "Rami" ; foaf:name "Rami Albatal" ;	100 1# a Albatal, Rami	100 1# a Albatal, Rami (entrada principal) ou 700 1# a Albatal, Rami (entrada secundária)
Identificador ORCID	<http://orcid.org/0000-0002-9269-8578> a prov:Person, foaf:Person ; rdfs:label "Rami Albatal" ; foaf:account <http://orcid.org/0000-0002-9269-8578#orcid-id> ;	700 17 a Albatal, Rami 0 (ORCID) 0000-0002-9269-8578 2 http://orcid.org/	100 1# a Albatal, Rami 0 (ORCID) 0000-0002-9269-8578 e/ou 700 1# a Albatal, Rami 0 (ORCID) 0000-0002-9269-8578
Dados biográficos	foaf:plan ""Data Scientist, Ph.D. graduated in 2010 from Grenoble University - France. Sing April 2012, I am the lead postdoctoral researcher the Lifelogging team at Insight Centre for Data Analytics - Dublin City University.	678 0# a Data Scientist, Ph.D. graduated in 2010 from Grenoble University - France. Sing April 2012, I am the lead postdoctoral researcher the Lifelogging team at Insight Centre for Data Analytics - Dublin City University.	N/A

XVIII ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO – ENANCIB 2017
23 a 27 de outubro de 2017 – Marília – SP

	My research focuses on three main area: Lifelogging, Automatic Image Classification and Machine Learning. After having achieved my Ph.D. and worked for two years on research related to visual content analysis and retrieval, I am currently working on proposing new generation of Lifelogging technologies that employ advanced sensed data analysis and machine learning, in the goal of building external human memory that supports the total recall and mining and improve data-derived decision making and planning. I am actively involved in organizing international conferences, reviewing high-level scientific journals, and seeking national, European and International funding opportunities for research projects." ;	My research focuses on three main area: Lifelogging, Automatic Image Classification and Machine Learning. After having achieved my Ph.D. and worked for two years on research related to visual content analysis and retrieval, I am currently working on proposing new generation of Lifelogging technologies that employ advanced sensed data analysis and machine learning, in the goal of building external human memory that supports the total recall and mining and improve data-derived decision making and planning. I am actively involved in organizing international conferences, reviewing high-level scientific journals, and seeking national, European and International funding opportunities for research projects.	
Lugar associado (local de nascimento)	< http://sws.geonames.org/2963597/ > rdfs:label "Ireland", "Éire" ; ns1:countryCode "IE" ; ns1:name "Ireland", "Éire" .	370 ## a Ireland 0 2963597 2 http://www.geonames.org	N/A
Fontes positivas de dados	foaf:page < http://ie.linkedin.com/in/ramialbatal >, < http://scholar.google.co.nz/citations?user=tuG6AIQAAAAJ&hl=en > ;	670 ## a http://ie.linkedin.com/in/ramialbatal 670 ## a http://scholar.google.co.nz/citations?user=tuG6AIQAAAAJ&hl=en	N/A

Fonte: Elaborado pelos autores - 2017.

Alguns dos dados coletados podem ser empregados para enriquecimento das autoridades, não tendo, obrigatoriamente, apresentação no catálogo bibliográfico. As informações coletadas poupam o tempo do bibliotecário ao dispensar pesquisas em fontes para identificação e estabelecimento da forma adotada para descrever um autor Pessoa. No exemplo coletado não constavam do RDF as informações de outras formas de identificação do autor (remissivas ver), que contribuiriam com a desambiguação de homônimos. Espera-se que estes dados sejam passíveis de coleta em extrações futuras.

Para este estudo foram considerados dados que não são voláteis, ou seja, que não são atualizados com frequência. Assim, o objetivo era avaliar a coleta de dados descritivos básicos, como o nome adotado, local de nascimento, fontes para informações e dados biográficos. Entretanto, pela oferta de dados disponíveis no ORCID, deve ser considerada no futuro, a possibilidade de reaproveitamento de outras informações, como, por exemplo, afiliação, local de trabalho, áreas de interesse e produções acadêmicas, com o intuito de serem coletadas de forma dinâmica, porém, com rotinas para atualização,

XVIII ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO – ENANCIB 2017
23 a 27 de outubro de 2017 – Marília – SP

de forma automática, sempre que ocorrerem alterações no perfil do pesquisador. Estes dados dinâmicos podem ser apresentados aos usuários diretamente no catálogo em linha, aumentando a oferta de informações para consulta.

Foram identificadas no formato MARC duas formas de representação do código ORCID para um autor Pessoa. Uma opção é utilizar a tag 024 no registro de autoridade:

024 7# |a 0000-0002-9269-8578 |2 ORCID

100 1# |a Albatal, Rami

Outra possibilidade é o uso da tag 700 da autoridade, conforme:

100 1# |a Albatal, Rami

700 17 |a Albatal, Rami |0 (ORCID) 0000-0002-9269-8578 |2 <http://orcid.org/>

Em ambos os casos, ao transpor o registro da autoridade ao registro bibliográfico, a representação do código ORCID será realizada na tag 100 |0, conforme:

100 1# |a Albatal, Rami |0 (ORCID) 0000-0002-9269-8578

A tag 100|a recebe a descrição de nomes de autoridades do tipo Pessoa, com o primeiro indicador definindo se a entrada foi feita pelo prenome (valor 0), sobrenome (valor 1) ou nome de família (valor 2). Outras tags podem ser incorporadas à descrição, com o intuito de proporcionar maior detalhamento da autoridade.

A tag 024|a acomoda um código identificador atribuído a uma autoridade. O primeiro indicador da tag 024 (valor 7) sinaliza que a fonte que originou o código será representada no subcampo |2 (ORCID). Assim, numa mesma tag é informado o código (024|a) e sua fonte identificadora (024|2). Este identificador, quando presente no registro da autoridade, proporciona mais elementos que corroboram a identificação única do autor que está descrito. Esta descrição resulta numa autoridade estabelecida como:

024 7# |a 0000-0002-9269-8578 |2 ORCID

100 1# |a Albatal, Rami

Já a tag 700 é utilizada para registrar cabeçalho que foi atribuído por outra instituição e que deseja-se vincular ao termo adotado na tag 100|a. No exemplo utilizado, a descrição foi feita pelo sobrenome do autor (Albatal, Rami), com inclusão da tag 700|a para sinalizar como outra instituição estabeleceu esta autoridade. No segundo indicador da tag 700 é informada de qual fonte a descrição foi retirada e, caso não corresponda à lista disponível (*Library of Congress, Medical Subject Headings, Canadian Subject Heading* etc.), seleciona-se o indicador 7, que permite o detalhamento da fonte utilizada no subcampo 700|2. No subcampo 700|0 é informado o código atribuído pela fonte identificada no campo 700|2, formando, assim, a descrição:

100 1# |a Albatal, Rami

700 17 |a Albatal, Rami |0 (ORCID) 0000-0002-9269-8578 |2 <http://orcid.org/>

**XVIII ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO – ENANCIB 2017
23 a 27 de outubro de 2017 – Marília – SP**

Neste estudo foi adotada a representação realizada por Godby, Wang e Mixter (2015), com descrição do código ORCID na tag 700 |0.

6 CONSIDERAÇÕES FINAIS

A catalogação tradicionalmente consiste na descrição das obras presentes em uma coleção por meio de aplicação de regras. Para favorecer a troca e intercâmbio de registros entre as instituições, os dados são representados em formato estruturado, onde o MARC 21 possui ampla utilização. Os princípios da catalogação foram definidos a partir do século XIX e sofreram poucas alterações até os conceitos estabelecidos pelo FRBR e FRAD, onde a descrição é realizada pela obra, com suas manifestações e expressões, e não somente pelo item, conforme é preconizado no AACR2. Este entendimento de representação está por alterar a forma como os dados de autoridades e bibliográficos são estruturados pelas bibliotecas, trazendo mudanças nas rotinas, adoção de outros formatos e ainda a possibilidade de explorar recursos de LD.

Com os catálogos disponíveis para consulta na Web e as facilidades para exploração de recursos da Web de dados, a possibilidade de utilizar aplicações da Web Semântica torna-se uma opção interessante por, além de facilitar o processo de catalogação ao reaproveitar dados existentes, permite a ampliação das possibilidades de identificação e relacionamento dos recursos do acervo com demais informações distribuídas na Web.

Os recursos do LD, além de contribuírem com a atividade de catalogação, também proporcionam enriquecimento do catálogo ao agregar informações complementares a autores e obras, favorecendo a contextualização ou ainda, proporcionando inferências e descoberta de recursos existentes, mas até então desconhecidos.

Conforme visualizado no Quadro 1 e no Quadro 1, os dados coletados no ORCID podem ser reutilizados no controle de autoridades. O ORCID, entretanto, contém somente dados de autores vinculados a instituições de pesquisa e ensino, não sendo aplicável aos demais autores. Assim, não atenderá a todas as autoridades. Outro ponto importante é a fase experimental na qual o ORCID se encontra, não permitindo pleno aproveitamento dos dados presentes no registro ao não exportá-los na totalidade para o modelo RDF.

Em análise inicial, além de identificar um autor de forma única, os dados oriundos do ORCID também permitirão aos usuários a descoberta de outras produções (*works*) realizadas por esta autoria e que não fazem parte do acervo da biblioteca, mas que podem ser localizadas em outros sítios, por meio de URIs.

As possibilidades de utilização de recursos do LD nos catálogos ainda não foram completamente avaliadas, mas representam uma transformação a partir do momento em que o catálogo deixa de ser um conjunto fechado de dados e passa a ter seus registros disponíveis e

XVIII ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO – ENANCIB 2017
23 a 27 de outubro de 2017 – Marília – SP

compartilhados na Web, podendo ter conexão com diversas outras bibliotecas e sítios. Os dados deixam de ser consumidos somente por pessoas, e passam a ser coletados e reutilizados também por máquinas. Com isso, o catálogo pode passar a ser o local que promove a reunião de informações, de fontes variadas, e as disponibiliza aos usuários, mesmo que os recursos não existam nos acervos. Assim o catálogo deixará de ser a representação das obras presentes em uma coleção e passa a ser um local de descoberta de diversos recursos, de acordo com os *datasets* com os quais o acervo está interligado.

REFERÊNCIAS

ARCHER, Phil. **Proposal for the improvement of the semantics of ORCID**s. 2015. Disponível em: <<https://www.w3.org/2015/03/orcid-semantics>>. Acesso em: 12 jul. 2017.

BARBOSA, A. P. **Novos rumos da catalogação**. Rio de Janeiro: Brasilart, 1978.

BERNERS-LEE, Tim. **Linked Data**. 2006. Disponível em: <<https://www.w3.org/DesignIssues/LinkedData.html>>. Acesso em: 13 jun. 2017.

BILDER, G. **Structure of the ORCID Identifier**. 2012. Disponível em: <https://docs.google.com/document/d/1awd6PPguRAdZsC6CKpFSSu1dullIT8E3kHwIJ3tD5o/edit>. Acesso em: 28 maio 2017.

BIZER, C.; HEATH, T.; BERNERS-LEE, T. Linked Data: the story so far. **International Journal on Semantic Web and Information Systems**, v.5, n.3, p. 1-22, 2009. Disponível em: <<https://eprints.soton.ac.uk/271285/1/bizer-heath-berners-lee-ijswis-linked-data.pdf>>. Acesso em: 26 jun. 2017.

CARTER, C. B.; BLANFORD, C. F. All authors must now supply ORCID identifiers. **Journal of Materials Science**, v. 52, n. 11, p. 6147-6149, Jun. 2017. Disponível em: doi:10.1007/s10853-017-0919-7. Acesso em: 28 jun. 2017.

CASSOTTA, M. L. J. et al. Recursos do conhecimento: colaboração, participação e compartilhamento de informação científica e acadêmica. **Informação & Sociedade: Estudos**, v. 27, n. 1, 2017. Disponível em: <<http://www.brapci.ufpr.br/brapci/v/a/23078>>. Acesso em: 28 maio 2017.

COYLE, K. **Library data in the web world**. Chicago: ALA TechSource, 2010. (Library Technology Reports ; n. 2). Disponível em: <<https://journals.ala.org/index.php/ltr/article/view/4673/5546>>. Acesso em: 12 jun. 2017.

DATAHUB. **ORCID. 2017**. Disponível em: <<https://datahub.io/dataset/orcid-dataset>>. Acesso em: 29 maio 2017.

DEMERANVILLE, Tom. **Using Schema.org for simple LOD representations within ORCID**. 2017. Disponível em: <<https://summit2017.lodlam.net/2017/06/23/using-schema-org-for-simple-lod-representations-within-orcid/>>. Acesso em: 12 jul. 2017.

GODBY, C. J.; WANG, S.; MIXTER, J. K. **Library linked data in the cloud: OCLC's experiments with new models of resource description**. Williston, VT: Morgan & Claypool, 2015. (Synthesis lectures on the semantic web: theory and technology; 9).

XVIII ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO – ENANCIB 2017
23 a 27 de outubro de 2017 – Marília – SP

HAAK, L. L. et al. ORCID: a system to uniquely identify researchers. **Learned Publishing**, v. 25, n. 4, p. 259-264, 2012. Disponível em: <<http://www.ingentaconnect.com/contentone/alpsp/lp/2012/00000025/00000004/art00004>>. Acesso em: 28 maio 2017.

HASTINGS, R. Linked Data in libraries: status and future direction. **Infoday.com**, Medford, MA, USA, v. 35, n. 9, nov. 2015. Disponível em: <<http://www.infoday.com/cilmag/nov15/Hastings--Linked-Data-in-Libraries.shtml>>. Acesso em: 16 jun. 2017.

INTERNATIONAL FEDERATION OF LIBRARY ASSOCIATIONS AND INSTITUTIONS (IFLA). **Declaração de Princípios Internacionais de Catalogação**. 2009. Disponível em: <http://www.ifla.org/files/assets/cataloguing/icp/icp_2009-pt.pdf>. Acesso em: 06 jun. 2017.

ISOTANI, S.; BITTENCOURT, Ig I. **Dados abertos conectados**. S. l.: Novatec, 2015. Disponível em: <<http://ceweb.br/livros/dados-abertos-conectados/>>. Acesso em: 24 jun. 2017.

LIBRARY OF CONGRESS. **Bibliographic framework as a Web of data**: Linked Data Model and supporting services. Washington, 2012. Disponível em: <<http://www.loc.gov/bibframe/pdf/marclid-report-11-21-2012.pdf>> Acesso em: 24 jun. 2017.

_____. **MARC 21 format for authority data**. 2017. Disponível em: <<http://www.loc.gov/marc/authority/>>. Acesso em: 06 jun. 2017.

LUBETZKY, S. **Principles of cataloging**. Oakland: Institute of Library Research University of California, 1969.

MEY, E. S. A.; SILVEIRA, N. C. **Catalogação no plural**. Brasília, DF: Briquet de Lemos, 2009.

OPEN KNOWLEDGE INTERNATIONAL. **The open data handbook**. 2004. Disponível em: <<http://opendatahandbook.org/guide/en/what-is-open-data/>>. Acesso em: 24 jun. 2017.

ORCID. 2017. Disponível em: <<https://orcid.org/content/about-orcid>>. Acesso em: 28 maio 2017.

_____. **Structure of the ORCID Identifier**. 2017b. Disponível em: <<https://support.orcid.org/knowledgebase/articles/116780-structure-of-the-orcid-identifier>>. Acesso em: 28 maio 2017.

RYAN, C. et al. Linked data authority records for Irish place names. **In J Digit Libr**, Berlin Heidelberg, v. 15, p. 73-85, abr. 2015.

SANTOS, P. L. V. A. da C.; PEREIRA, A. M. **Catalogação**: breve história e contemporaneidade. Niterói: Intertexto, 2014.

W3C. **SPARQL: 1.1 visão geral**. 2013. Disponível em: <<https://www.w3.org/TR/sparql11-overview/>>. Acesso em: 24 jun. 2017.

WALTON, H. ORCID: integration among publishing and funding organizations: an examination of process and rationale. **School of Information and Library Science**. 2014. Disponível em: <<https://cdr.lib.unc.edu/record/uuid:d60e4bb7-d4d0-49a5-9e84-7bde3952e838>>. Acesso em: 29 maio 2017.

WILSON, B.; FENNER, M. Open Researcher & Contributor ID (ORCID): Solving the Name Ambiguity Problem. **Educause Review**, S.l., v. 47, n. 3, p.54-55, maio 2012. Disponível em:

XVIII ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO – ENANCIB 2017
23 a 27 de outubro de 2017 – Marília – SP

<<http://er.educause.edu/articles/2012/5/open-researcher--contributor-id-orcid-solving-the-name-ambiguity-problem>>. Acesso em: 1 jun. 2017.