

XIX encontro nacional
de pesquisa em
ENANCIB ciência da informação

// SUJEITO INFORMACIONAL E AS
PERSPECTIVAS ATUAIS EM CIÊNCIA
DA INFORMAÇÃO. //

22-26
OUTUBRO
2018
LONDRINA/PR



XIX ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO – ENANCIB 2018

GT-8 – Informação e Tecnologia

ONTOLOGIAS SPAR E USO DE METADADOS FRBR PARA ELABORAÇÃO DE REFERÊNCIAS

Marcos Luiz Mucheroni (Universidade de São Paulo – USP)

José Fernando Modesto da Silva (Universidade de São Paulo – USP)

Erika Alves dos Santos (Universidade de São Paulo – USP / Fundação Jorge Duprat Figueiredo de Segurança e Medicina do Trabalho – Fundacentro)

SparOntologies and FRBR metadata usage for references elaboration

Modalidade da Apresentação: Comunicação Oral

Resumo: Trata de estudo descritivo e exploratório, que integra pesquisa de doutorado em Ciência da Informação, na qual são problematizadas as normas de referências com uso das novas tecnologias. Discute-se a normalização de referências no contexto Brasileiro, a partir da análise da estrutura de apresentação da norma ABNT NBR 6023 – Informação e documentação: Referências – elaboração, considerando inclusive, aspectos de seu projeto de revisão. A partir da análise desta norma, que identificou diversas deficiências no contexto da normalização de referências no Brasil, a discussão avança para o contexto mundial onde o uso das referências como faceta da representação descritiva já é feito com uso de tecnologias, em especial as ontologias SPAR, especificamente FaBiO (*FRBR aligned Bibliographic Ontology*) e CiTO (*Citation Typing Ontologies*), com destaque para o modelo de relações de metadados proposto pelo FRBR. Sugere o estabelecimento de interoperabilidade, em bancos de dados, de metadados entre o VIAF e OpenCitations, para simplificar o processo de elaboração de referências, do ponto de vista dos pesquisadores e, estabelecer um formato único de representação da informação na forma de referências, de modo a consolidar o princípio da normalização e favorecer a recuperação da informação, isto significa uma forma final de especificação entre a Web Semântica e as exigências de referências normalizadas.

Palavras-chave: Referências; FRBR; SPAR; VIAF; OpenCitations.

Abstract: This paper was derived from a doctoral thesis where the norms of references with the use of new technologies are problematized. It discusses the normalization of references in the Brazilian context, based on the analysis of the structure of presentation of the standard ABNT NBR 6023 - Information and documentation: References - elaboration. Based on this analysis, which identified

several deficiencies in the context of the Brazilian normalization references, the discussion advances into the global context where there is use of references as a facet of previous descriptive representation, using technologies, especially SPAR and FaBiO (FRBR aligned Bibliographic Ontology) and CiTo (Citation Typing Ontologies). Since this ontology has metadata relations used by the FRBR, it proposes the interoperability, in databases, of VIAF and OpenCitations, to simplify the process of making references from the researchers point of view and to establish a unique format of information representation in references, in order to consolidate the principle of normalization and favor information retrieval. Resulting in a final form of specification between the Semantic Web and the requirements of standardized references.

Keywords: References; FRBR; SPAR; VIAF; OpenCitations.

1. INTRODUÇÃO

A demanda atual na organização da informação nem sempre conserva aspectos cognitivos essenciais ao processo de conversão de dados em informação, como é o caso dos fundamentos da análise documentária que, na prática, se convencionou chamar de linguagens documentárias, enquanto instrumentos para a organização da informação, que configuram uma negligência em tempos de expansão tecnológica, pois, ambas constituem os pilares da recuperação da informação.

Os recursos tecnológicos, especificamente a Web Semântica, atribuíram um caráter relativista a tal consideração, uma vez que há contextos em que o tratamento adequado da informação além de identificar, também promove a interação e informações sobre a localização da informação estabelecendo, inclusive, relações com conteúdos correlatos, conforme propõem as ontologias *Semantic Publishing and Referencing* (SPAR Ontologies).

Os oito anos de existência das Ontologias SPAR, foram marcados por diversos desenvolvimentos e aprimoramentos, que ainda não atingiram sua completude e demandam revisão (PERONI, 2018) mas que, apesar disso, convergem para um triunfalismo de referências feitas automaticamente em linguagens naturais (PERONI, 2017).

As técnicas contemporâneas de gestão da informação na Web para a (co)decodificação da informação da informação possibilitam a consulta, a integração e a correlação de metadados provenientes de múltiplas fontes de forma automatizada. As ontologias SPAR, especificamente a FRBR *Aligned Bibliographic Ontology* (FaBiO) e *Citations Typing Ontologies* (CiTO) que complementam, aprimoram e convergem para as descrições FRBR, já são aplicadas e desempenham um papel fundamental nesse processo, porém, é patente a necessidade de que tais conceitos sejam incorporados pelos editores como uma forma de refletí-los na comunicação científica representada nos artigos, livros e anais de eventos, como forma de

unificar os interesses editoriais, autorais e acadêmico-científicos, visando, dentre outras coisas, a revolução nas formas de publicação semântica (PERONI; SHOTTON, 2012).

A citação bibliográfica, ou seja, o ato de referir-se de uma entidade citada, é uma das atividades mais importantes de um autor na produção de uma obra bibliográfica, já que o reconhecimento de fontes que essa atividade representa se situa no cerne da atuação acadêmica. A rede de citações criada pela combinação de informações de citações de artigos acadêmicos e livros é uma rica fonte de informações para os acadêmicos e pode ser usada pelos editores para criar formas de navegar em seus dados, bem como para calcular métricas que reflitam a importância de um periódico (por exemplo, o fator de impacto) ou um autor (por exemplo, o índice H) (PERONI, SHOTTON, 2012).

Todavia, o tratamento da informação é uma tarefa de caráter não exclusivo aos profissionais de Ciência da Informação, considerando a publicação científica como o principal meio de comunicação de achados científicos. E, em sendo assim, a representação descritiva, na forma de referências, desempenha uma função relevante nos aspectos de encontrabilidade da informação. Mesmo diante do uso ascendente dos instrumentos automatizados para a redação de referências, tal atividade normalmente é conduzida pelos próprios pesquisadores nas diversas áreas do conhecimento. Isto reforça a importância de que os padrões de apresentação de metadados para a elaboração de referências se apresentem em linguagem acessível ao senso comum e de fácil entendimento por pesquisadores, profissionais e estudantes, que nem sempre dominam todos os termos técnicos específicos da Biblioteconomia.

2. O TRATAMENTO DA INFORMAÇÃO SOB A ÓTICA DAS ONTOLOGIAS

O tratamento destinado à informação é determinante para a contextualização e atribuição de valor a esta e, deve ser considerado sob duas vertentes: a primeira, sob a visão de quem procede ao tratamento da informação e a segunda, sob a visão de quem a recebe e determinará o seu uso ou descarte, conforme a pertinência e relevância para o trabalho do pesquisador. Partindo do ponto de vista de quem procede ao tratamento da informação, há ainda, duas ramificações: uma conduzida por profissionais da Ciência da Informação e outra, conduzida por pesquisadores de outras áreas do conhecimento que, geralmente procedem ao tratamento primário da informação, considerando como tal algumas atividades inerentes

à pesquisa como, por exemplo, a elaboração de resumos e abstracts, além da atribuição de palavras chave e redação de referências.

Tal argumento sugere reflexões sobre a adequação do contexto de determinados termos, na sua produção, no armazenamento, na busca, na recuperação e no uso da informação que representam, considerando suas possíveis variações de acordo com a perspectiva abordada. Uma significação que eventualmente possa ser considerada “universal”, nesse contexto, seja o entendimento de “tratamento adequado da informação” como aquele em que ao leitor (pesquisador) são oferecidos todos os elementos necessários à identificação, individualização, localização e contextualização da informação representada, de modo a evidenciar as abordagens contidas em um determinado documento.

Cerne da comunicação científica, a representação descritiva, tradicionalmente reconhecida na forma da catalogação, foi moldada pelo cenário da organização da informação que demandou (e continua demandando) adaptações na gestão e tratamento da informação. Nesse novo contexto, a interoperabilidade deixou de ser característica e assumiu o caráter de requisito, sobretudo considerando a apresentação multiforma da representação descritiva. Analogamente, a atribuição de significância semântica a um documento lhe confere o destaque de seu significado, eleva seus índices de encontrabilidade, possibilita sua ligação semântica a outros documentos relacionados, favorece o acesso à informações correlatas de forma ativa e automática e promove a integração entre os documentos. Tais fatores justificam a necessidade de desenvolvimento de modelos semânticos (vocabulários e ontologias) e respectiva documentação que as torne compreensíveis a editores e à comunidade científica de modo geral, além de novos algoritmos para extrair da tecnologia o máximo possível em favor da gestão, organização e encontrabilidade da informação. Nesse contexto, as ontologias assumem o papel e a função de atribuir características conceituais e semânticas aos metadados, garantindo-lhes a permeabilidade da informação entre distintos sistemas automatizados (SHOTTON, 2012).

Já há trabalhos na CI (BERTI JUNIOR *et al.*, 2017), que já propõe sistemas semiautomáticos de indexação, e que contribuem para compreensão do tema.

A Web Semântica e o uso de ontologias na atribuição de significados ao registro de metadados constituem aspectos que há alguns anos ocupam postos de destaque na Ciência da Informação e se materializam nas ontologias utilizadas na descrição bibliográfica, como as *Bibliographic Ontology*(BiBo), o vocabulário *Dublin Core* (DC) e as Ontologias SPAR, sendo

essas últimas as mais recentes e mais completas para a descrição bibliográfica em ambiente digital.

3. AS ONTOLOGIAS SPAR

Com origens de natureza filosófica, as ontologias podem ter diferentes significados, dependendo da área de abordagem. Na Ciência da Informação e na Ciência da Computação, cujas atuações se entrelaçam cada vez mais, as ontologias são os instrumentos que permitem a nomeação formal e definição de categorias, propriedades e relações entre conceitos, entidades e dados atrelados a um ou mais domínios ou objetos de informação. Em uma correlação à Ciência da Informação tradicionalista, as ontologias podem ser comparadas, a grosso modo, como uma releitura da indexação (e toda a documentação que a fundamenta), voltada para o universo semântico digital.

As Ontologias SPAR são um conjunto de ontologias complementares e ortogonais da *Web Ontology Language* (OWL 2 DL), para a estruturação de metadados bibliográficos legíveis por máquina em formato RDF, que primam pela descrição de todos os aspectos editoriais e referenciais de uma publicação: descrição do documento, identificadores de recursos bibliográficos, tipos de citações e contextos relacionados, referências bibliográficas, partes e status de documentos, papéis e contribuições de entidades, dados bibliométricos e processos de fluxo de trabalho (SANTOS; MUCHERONI, 2018).

O conjunto, ainda em desenvolvimento, integra 8 ontologias: *The Bibliographic Reference Ontology* (BiRO); *The Citation Counting and Context Characterization Ontology* (CAO); *The Document Components Ontology* (DoCO); *The Publishing Roles Ontology* (PRO); *The Publishing Status Ontology* (PSO); *The Publishing Workflow Ontology* (PWO); *The FRBR-aligned Bibliographic Ontology* (FaBiO) e; *The Citation Typing Ontology* (CiTO), sendo as duas últimas avocadas como objeto desta reflexão.

A metodologia proposta envolve, portanto, ontologias e os bancos de dados usados.

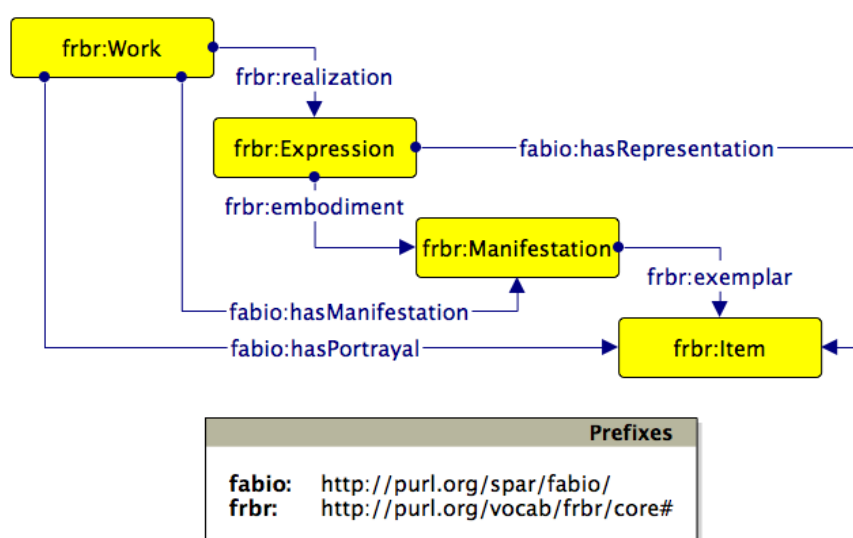
3.1. *The frbr-aligned bibliographic ontology* (fabio)

A partir das mudanças nas formas de produção, armazenamento, busca, recuperação, acesso e uso da informação ao longo das décadas, se desencadeou um processo (necessário) de reinvenção da Ciência da Informação (SOERGEL, 2001), por meio do qual se deu a adequação de conceitos e práticas que sugerem o vislumbre de um longo caminho ainda a ser percorrido. Uma dessas mudanças foi a remodelagem da representação descritiva, que

dentre outras coisas, introduziu o modelo conceitual de descrição bibliográfica *Functional Requirements for Bibliographic Description* (FRBR) no qual se fundamenta a estrutura das classes FaBiO.

A ontologia FaBiO amplia as possibilidades de relacionamentos entre as entidades envolvidas em um recurso de informação propostas pelo FRBR, de modo a permitir não apenas as relações entre obras e expressões, expressões e manifestações e, manifestações e itens, mas estendendo as possibilidades de relacionamentos entre todas as entidades envolvidas em uma obra, conforme ilustra a figura 1 – Propriedades e relações adicionais FRBR segundo a ontologia FaBiO. Além disso, como compromisso com a interoperabilidade, a ontologia FaBiO importa entidades de outras ontologias como por exemplo, *Dublin Core Terms*, *PRISM* e *SKOS*.

Figura 1: Propriedades e relações adicionais FRBR segundo a ontologia FaBiO.



Fonte: <http://www.sparontologies.net/ontologies/fabio> Adaptado de Peroni (2018)

A ontologia FaBiO está alinhada com os parâmetros conceituais apresentados pelo FRBR, contudo, as necessidades de adequação às mudanças do universo da informação fomentaram a atualização deste modelo conceitual, que culminaram na origem do *IFLA Library Reference Model* (IFLA-LRM), publicado em agosto de 2017. Estabeleceu-se aí o “efeito cascata”: o IFLA-LRM impactou no *Resource Description and Access* (RDA), que substituiu o FRBR pelo IFLA-LRM como modelo conceitual basal (RDA STEERING COMMITTEE, 2017). A apreciação lógica deste cenário sugere a possibilidade de que a ontologia FaBiO, que está alinhada ao modelo FRBR, seja remodelada nos termos do novo modelo conceitual IFLA-LRM.

3.2. *The Citation Typing Ontology (CiTO)*

Denomina-se citação a menção de outros documentos, ou trechos destes, em trabalhos científicos e, seu registro se dá por diversas razões: endosso, contestação ou complementação de uma argumentação, exemplificação, contextualização, definição, dentre outros, a ontologia CiTO tem por objetivo geral de contemplar estas necessidades.

Enquanto a tecnologia, com seus sistemas de publicação online e instrumentos de busca, tornou factível a encontrabilidade de trabalhos de forma individual, a ontologia CiTO evidencia a rede de conhecimento que existe na literatura científica, por meio das relações entre trabalhos, autores e projetos de pesquisa. A ontologia CiTO orienta a geração de metadados para categorização do tipo e natureza das citações, diretas e indiretas, nos aspectos factuais e retóricos (SHOTTON, 2010b). Isto permite que os relacionamentos entre as publicações sejam estabelecidos não apenas pelas entidades envolvidas, mas também pelo seu conteúdo, considerando que as citações, conjuntamente com as referências, constituem fontes de informação, enquanto meios de acesso para um agrupamento de documentos pré-selecionados com temática relacionada. Assim, o estabelecimento de uma rede de conhecimento, baseada em citações; a tipificação das citações em um formato legível por máquina; a quantificação da frequência destas e, a caracterização do trabalho citado são os principais objetivos da ontologia CiTO. Pela própria natureza de seu conteúdo, documentos de natureza artística, não são passíveis de descrição com a ontologia CiTO, justamente por não incluírem citações. Por tal motivo, na ontologia CiTO obra é uma subclasse do conceito obra adotado pelo modelo FRBR (SHOTTON, 2010a).

4. A ELABORAÇÃO DE REFERÊNCIAS SEGUNDO A ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS (ABNT)

As citações constituem um dos requisitos para a atribuição do caráter científico a uma publicação. Contudo, estas não se sustentam por si, de modo que a inclusão de uma citação em um determinado trabalho implica, obrigatoriamente, na indicação da respectiva referência, independente do estilo bibliográfico utilizado. Em contrapartida, a preocupação com o tratamento desta forma de representação da informação não é evidente, sobretudo no ambiente semântico, como se a padronização na descrição de metadados fosse o único elemento determinante para a encontrabilidade da informação representada. Enquanto há o empenho na atualização e adequação dos elementos norteadores da organização da informação, a exemplo dos modelos conceituais de representação da informação (FRBR, RDA

e outros) e das ontologias (SPAR ontologies, OWL, SKOS, etc.), face aos novos rumos da Ciência da Informação e da gestão e organização do conhecimento, não se percebe a empenhos convergentes no âmbito da elaboração, gestão e uso da informação representada na forma de referências, pelo menos no contexto brasileiro.

Tal cenário sugeriu considerações críticas sobre a norma ABNT NBR 6023 – Informação e documentação: Referências: elaboração, publicada pela Associação Brasileira de Normas Técnicas (ABNT). A norma, amplamente adotada no Brasil e cuja edição mais recente data de 2002 é baseada nas normas ISO 690:1987 e ISO 690-2:1997, publicadas pela *International Standardization Organization (ISO)*, ambas canceladas e substituídas pela norma ISO 690:2010 em 14 de junho de 2010. Tais considerações fundamentam-se nos resultados obtidos a partir de análise crítica de conteúdo das normas supra.

As análises sugeriram que a norma ABNT NBR 6023 – Informação e documentação: Referências: elaboração é deficitária no cumprimento de seu propósito de fornecer subsídios para a redação de referências de forma clara, concisa e, principalmente, padronizada. A norma não instrui sobre os procedimentos descritivos adequados para alguns tipos de documentos, a exemplo das entrevistas. Isto favorece a ocorrência de redações de referências distintas para um mesmo documento. Também inexistem indicações de procedimentos descritivos para documentos nos quais não consta data de publicação. Também são deficitárias as orientações normativas sobre a ordenação da lista de referências: não há menção procedimental para a ordenação correta de duas ou mais referências de documentos de mesma autoria e sem indicação de data de publicação, nem tampouco procedimentos para referenciar partes distintas, porém não seguidas e de mesma autoria, na mesma obra.

Estas e outras questões concorrem para a não completude do alcance dos objetivos estabelecidos pela ABNT NBR 6023:2002, no sentido de “orientar a preparação e compilação de referências do material utilizado para a produção de documentos e para a inclusão de bibliografias, resumos, resenhas, resenhas, resenhas e outros.” (ABNT, 2002, p.1).

Além disso, a norma ABNT NBR 6023:2002 determina, em caráter obrigatório em alguns casos, a descrição de metadados dispensáveis para a recuperação da informação. Um exemplo é a indicação da cidade de publicação de publicações periódicas. Ora, o local de publicação de periódicos é uma informação que nem sempre é tratada com destaque pelos editores e pode, inclusive, demandar pesquisas em fontes externas ao documento em si. Tal obrigatoriedade, além de gerar dificuldades para o pesquisador, que nem sempre sabe onde

localizar tais informações, demonstra que a norma ABNT NBR 6023:2002 não está alinhada com os recursos tecnológicos para a recuperação da informação. Outro sinal de sua obsolescência é o fato de que dos 4 documentos prescritivos à referida norma, pelo menos 1 está obsoleto: a norma ABNT NBR 10522:1988 Abreviação na descrição bibliográfica, foi cancelada, sem substituição, em 30 de setembro de 2003.

As normas ISO 690:1987 e ISO 690-2:1997, nas quais está baseada a ABNT NBR 6023:2002, foram ambas substituídas pela norma ISO 690:2010. Outro instrumento prescritivo é o Código de Catalogação Anglo Americano (CCAA2), (em inglês, *Anglo American Cataloging Rules – AACR2*), que tende a ser gradativamente substituído pelo *Resource Description and Access (RDA)*. Conste, entretanto, que a ABNT NBR 6023:2002 permaneceu em processo de revisão de março de 2009 até março de 2017. O primeiro projeto de revisão desta norma circulou em consulta nacional entre agosto e outubro de 2015, enquanto o segundo projeto de revisão, que data de junho de 2018, foi disponibilizado para consulta pública em julho de 2018.

Embora tal projeto de revisão não configure o ponto focal desta reflexão, algumas observações merecem destaque: as omissões de informações nas orientações descritivas tendem a persistir, em face da adoção frequente de termos e expressões genéricas como “entre outros”. É compreensível que em alguns casos, o detalhamento da informação se torna inviável pela sua extensão, porém, deixar de fazê-lo pode representar óbices aos objetivos normativos de alcançar padrões mínimos de uniformidade em suas áreas de competência, na medida em que tais omissões dão abertura para interpretações distintas, e até opostas em diferentes contextos, que podem culminar em aplicações múltiplas de uma mesma instrução normativa. A segunda revisão da norma também não esclarece a diferença entre bibliografia e referência. Outra aparente novidade é a possível previsão de extração de metadados a partir de outras fontes de informações além do próprio documento, o que já é uma tendência na representação descritiva, analogamente às prescrições do RDA. Em contrapartida, o projeto de revisão retirou a indicação dos códigos de catalogação como documentos prescritivos, o que pode caracterizar um retrocesso, ao dissipar a aproximação entre a catalogação e a elaboração de referências, enquanto facetas da representação descritiva. Também pode ser alvo de questionamento a efetividade de uma norma desenvolvida (ou revisada) em um cenário no auge de um processo de redesenho na organização da informação, tal como ocorre

atualmente com a representação descritiva e, sobretudo, considerando a notável irreflexão de tais conceitos no conteúdo da norma.

Cumulativamente, tais pontos fomentam múltiplas interpretações das diretrizes para a elaboração de referências, que culminam no aumento da probabilidade de ocorrência de referências distintas para um mesmo documento. As deficiências identificadas na representação da informação na forma de referências no cenário brasileiro são passíveis de ocorrer também em outros países, tendo como agravante a existência de cerca de 8 centenas de estilos bibliográficos distintos. Uma vez que pesquisadores não se limitam a publicar trabalhos apenas em seus países de origem, o conhecimento sobre as diretrizes estabelecidas pelos estilos bibliográficos vigentes em cada país (ou área do conhecimento) para os quais venham a submeter suas contribuições se torna um requisito.

Em resposta a tal cenário, alguns editores científicos disponibilizam instrumentos de elaboração automática de referências, baseados nos inúmeros estilos bibliográficos vigentes pelo mundo. Tais instrumentos demandam o preenchimento, de forma prévia e manual, dos campos de metadados correspondentes à publicação a ser referenciada, os quais são armazenados apenas localmente. Não há uma base de dados para armazenamento compartilhado e reutilização destes metadados, de modo que a cada vez que um mesmo documento é referenciado, os campos de metadados devem ser novamente preenchidos, em contradição aos princípios do compartilhamento de dados e da interoperabilidade. Desconsiderem-se no âmbito dessa discussão o tempo necessário para adequação desses softwares às revisões das normas de referências em todo o mundo e, também, a fundamentação conceitual e os documentos prescritivos de cada uma das normas de referências vigentes no mundo, sob o risco de adentrar questionamentos que extrapolam a delimitação desta discussão. Porém, se as referências são uma faceta da representação descritiva e, se as normas são instrumentos delimitadores de suas formas de apresentação, visando, sobretudo, a padronização, a variância no conjunto documental que fundamenta os estilos bibliográficos é questionável e contradiz os princípios da normalização.

5. A PROBLEMATIZAÇÃO DA NORMALIZAÇÃO E MANIPULAÇÃO DE REFERÊNCIAS E METADADOS

Como instrumentos de controle e acesso à informação, catálogos bibliográficos e referências são abrangidos pelo conceito de representação descritiva e se constroem a partir de diretrizes estabelecidas por padrões e normas de amplo espectro de reconhecimento e

uso. Contudo, o conjunto documental que norteia a representação descritiva, especialmente as diretrizes para a elaboração de referências, se apresenta segundo a linguagem específica da Ciência da Informação com todo o seu arsenal de abreviaturas, siglas e conceitos, mesmo em alguns dos padrões bibliográficos de maior utilização. Inclusive, há estilos bibliográficos que apontam as diretrizes de descrição documentária determinadas pelo *Anglo American Cataloging Rules* (AACR2) como instrumentos prescritivos e que recomendam a sua consulta para o fomento de decisões mediante situações não abrangidas pelo próprio estilo bibliográfico. É o caso da norma ABNT NBR 6023:2002. Tal cenário evidencia a fragilidade na interpretação e uso da norma em pauta, que se revela como uma das possíveis causas para a elaboração de padrões próprios, ou versões interpretadas dos estilos bibliográficos pelos editores científicos e instituições de ensino. Sobre a existência de diversos estilos bibliográficos e suas múltiplas interpretações, Bertholino e Silva (2008, p.40), pontuam que:

[...] há uma miscelânea de tipos de apresentação. Os artigos encontrados nas revistas apresentam padrões que em determinadas situações seguem interpretações dos editores.

Um exemplo eminente é a adoção, pela área de Ciências da Saúde ou Biomédicas, da norma Vancouver que difere em forma de apresentação das referências preconizadas pela ABNT.

Surgem, então, questionamentos e dúvidas e uma grande confusão de interpretação pelos alunos e/ou autores que querem padronizar sua produção científica.

[...]

Seria extremamente interessante que algum tipo de consenso pudesse ser alcançado a fim de que as comunidades científicas, acadêmicas e editores trabalhassem como uma única norma evitando as disparidades, hoje observadas, na normalização.

Apesar da factibilidade na correspondência entre catalogação e elaboração de referências, a restrição de sua materialização apenas por meio dos instrumentos documentais de apresentação das diretrizes, em ambas as atividades, pode representar óbices para o entendimento da forma adequada de apresentação de metadados em citações e referências, principalmente por pesquisadores de outras áreas que não a Ciência da Informação. Em contrapartida, a normalização requer noções de documentação nem sempre familiares aos pesquisadores.

Em uma realidade em que, cada vez mais se tende a manipular grandes massas de dados, dentro do princípio cibernético de que todo esforço a mais

na entrada do processo (aqui visto amplamente, como a geração do conhecimento) redundando em uma esperada economia na saída (recuperação e uso da informação), a normalização surge como fator não só de qualidade, mas como facilitador da transferência da informação científica. (RODRIGUES; LIMA, GARCIA, 1998, p.155).

A codificação da informação em representações passíveis de processamento automático é um dos meios que permitem que as máquinas façam o processamento que atualmente, na web sintática, exige a intervenção humana. O processamento automático da informação pode ser um grande aliado no sentido de prover informações relevantes no tempo e espaço correto e, pode trazer benefícios em diversos aspectos, desde o comércio eletrônico, até o gerenciamento do conhecimento e sua aplicação pelo usuário final, por meio da indexação de informações em mecanismos de busca e da utilização de agentes de softwares que agem de maneira independente e têm autonomia para atuar em nome dos interesses de seus usuários. (BREITMAN, 2014). Isto reforça a necessidade de repensar as formas de representação da informação que a partir da aplicação de funcionalidades semânticas, considerando que pode estabelecer uma aproximação exitosa entre a catalogação e a apresentação de referências, no sentido de agregar informações de forma independente e proativa.

O que é necessário é uma forma de representar o conhecimento que permita aos computadores tanto interpretá-lo no sentido tradicional de exibi-lo na tela em um formato legível por humanos, quanto compreendê-lo em um nível de computador, permitindo assim que o computador reaja de forma autônoma a esse conhecimento. (KÜCK, 2004, tradução nossa).

Neste sentido, o compartilhamento e intercâmbio de metadados entre VIAF e Open Citations representa a materialização do modelo entidade-relacionamento proposto pelo FRBR, e pode trazer benefícios ao contexto da normalização de referências, seja pelo controle das formas de registro das autoridades, seja pela atribuição do caráter inteligível e semântico às referências pelas máquinas ou, pelo desencargo da comunidade científica em interpretar as diretrizes normativas estipuladas por cada estilo bibliográfico. O trabalho cooperativo e o intercâmbio de metadados são iniciativas consagradas na Ciência da Informação, porém, é admirável o fato de que embora inúmeros estudos demonstrem a efetividade e os benefícios oriundos da aplicação dessas filosofias de trabalho para a organização e encontrabilidade da informação, parece não haver movimentação efetiva para a aplicação de tais técnicas em favorecimento da representação da informação na forma de referências. Tal fato, somado à complexidade envolvida no processo da estruturação de metadados, denuncia uma postura

(involuntária) da Ciência da Informação que além de depreciar e contradizer o propósito da organização e recuperação da informação, ainda sugere que o compromisso com o registro da informação é uma atribuição exclusiva dos bibliotecários, até mesmo porque, os instrumentos reguladores dessa atividade, tal como as normas e guias de utilização de estilos bibliográficos se apresentam e em uma linguagem técnica, vulgarmente chamada de “biblioteconomês”, que dificulta a interpretação de pesquisadores de áreas além da competência da Ciência da Informação. Contudo, não se pode ignorar que a redação de referências integra ativamente o cenário da organização e recuperação da informação, em alguns casos de forma mais efetiva do que os próprios catálogos bibliográficos e que, a sua redação e uso extrapola a atuação profissional do bibliotecário, sendo exercida por grande parte dos pesquisadores.

O compartilhamento de metadados para a elaboração de referências pode ser uma alternativa para tornar mais efetiva a normalização na redação de referências. Inclusive, havendo um repositório compartilhado de acesso aberto no qual as referências possam ser recuperadas, extingue-se a necessidade da existência de estilos bibliográficos múltiplos em prol da garantia da uniformidade na apresentação de referências. De fato, a forma de registro de metadados é determinante para a recuperação, ou não, de uma informação codificada.

[...] a relativa falta de experiência de muitos acadêmicos neste novo contexto internacional propicia alguns problemas que afetam a visibilidade de seus trabalhos. A isto há que se somar as dificuldades decorrentes das diferenças culturais entre os criadores das principais bases de dados [...] e os diversos contextos nacionais aos que se projetam como expressão de sua globalização. Produz-se assim uma grande quantidade de desencontros e erros que afetam a qualidade e eficácia dos processos de codificação e recuperação da informação.

[...] Os problemas com a correta identificação dos autores preocupam não somente aos envolvidos, mas também às grandes empresas que gerenciam as principais bases de dados bibliográficos que também pretendem reduzir o número de erros e desenvolver ferramentas que permitam buscas mais precisas (ALIAGA; CORREA, 2011, tradução nossa).

O uso de ontologias como forma de auxílio a organização da informação, não apenas tem um desenvolvimento emergente, como já há resultados (BERTI JUNIOR et al., 2017), se dirigida a um domínio.

6. VIAF E OPENCITATIONS COMO CANAIS DE GESTÃO E INTERCÂMBIO DE METADADOS

Duas iniciativas promissoras nesse contexto são o VIAF e o OpenCitations. O primeiro, trata de um consórcio internacional do qual participam mais de 50 instituições, na sua maioria bibliotecas nacionais, que visa estabelecer entradas padronizadas e o intercâmbio de dados de autoridades para entidades pessoais, corporativas e eventos, bem como o relacionamento com suas respectivas obras. O OpenCitations, por sua vez trata de um repositório de acesso aberto para o compartilhamento de metadados referentes a citações e referências.

A interoperabilidade entre as duas iniciativas, VIAF e OpenCitations, por meio de um modelo RDF é potencialmente favorável à universalização das formas de registro de referências e ao desengargo da comunidade científica no que se refere ao manuseio de múltiplos estilos bibliográficos, considerando que a interpretação dos instrumentos normativos relacionados à representação descritiva nem sempre é uma tarefa simples e/ou desempenhada de forma adequada por pesquisadores alheios à Ciência da Informação.

O relacionamento e o compartilhamento de metadados entre VIAF e OpenCitations, além de representar uma possível alternativa para a simplificação e unificação das formas de redação e apresentação de referências, uma vez consolidado, também tornaria injustificável a existência de centenas de estilos bibliográficos distintos. Ter-se-ia então consolidada a normalização na representação da informação em forma de referências, que poderiam inclusive, constar nos documentos na forma de um *link*, que remeteria o consulente para um ambiente digital externo, onde as informações poderiam ser descritas na sua forma completa inclusive, com possibilidade de acesso ao documento na íntegra, em alguns casos. Incumbiria ao pesquisador, portanto, apenas o registro do link correspondente ao recurso que contivesse os metadados correspondentes à publicação referenciada. A redação e registro da referência, em contrapartida, deveria ocorrer uma única vez, sob a responsabilidade de profissionais da Ciência da Informação, tal como já ocorre na filosofia do trabalho cooperativo há anos consolidado na área.

Além de conferir maior visibilidade às publicações, por meio da uniformidade na apresentação de referências, o tratamento semântico da informação contida nos instrumentos de compartilhamento de metadados, VIAF e OpenCitations, por meio da atribuição de contextualização e significado proporcionado pelas Ontologias SPAR, potencializaria e promoveria o uso de ambos os instrumentos além de simples fontes de consulta para o estabelecimento de entradas e cabeçalhos autorizados ou formas padronizadas de descrição de citações e referências.

O estabelecimento de relações entre as entidades de qualquer natureza, consolida uma rede de informação intangível, porém proporcionalmente poderosa no sentido de reunir informações correlatas e, nesse sentido, tais iniciativas poderiam ser consideradas como elementos de ampliação das funções do VIAF e OpenCitations enquanto plataformas de relacionamento de metadados, com efetiva contribuição positiva para a recuperação da informação e, tampouco, para a apresentação de citações e referências nos documentos científicos, de forma adequada.

Nesse contexto, a atribuição de significados, o compartilhamento e a ligação de dados compõem a tríade (sujeito, predicado e objeto), que rege a organização da informação na era pós-semântica e consolida as relações interdisciplinares entre a Ciência da Informação e a Informática. Entende-se que a aplicação das tecnologias semânticas, pautadas nos princípios de relações de metadados de que trata o FRBR, às referências, além de lhes agregar valor enquanto fontes de informações, também tornaria injustificável a multiplicidade de estilos bibliográficos. Assim, a referência assumiria efetivamente o papel de fonte de informação, como uma extensão do catálogo bibliográfico.

7. CONSIDERAÇÕES FINAIS

As contribuições da tecnologia para a Ciência da Informação no que se refere à organização e gestão da informação são cada vez mais presentes e evidentes. Entretanto, não há evidências de uma discussão organizada sobre a utilização de dispositivos semânticos para a gestão e recuperação da informação. Embora as tecnologias tenham conquistado popularidade e reconhecimento social, ainda há muito que se avançar no âmbito da Ciência da Informação. A indústria do comércio eletrônico e das mídias de redes sociais está explorando cada vez mais as funcionalidades semânticas e de dados ligados (linked data) (BIZER; HEATH; BERNERS-LEE, 2009), ironicamente, de forma mais evidente e eficaz do que a Ciência da Informação, em alguns aspectos, mas há trabalhos emergentes na área.

No que se refere à recuperação da informação por meio de referências, a simples supressão ou inclusão de instruções nas normas pode tornar o cenário da gestão da informação confuso para a maioria dos pesquisadores, a quem a complexidade do tema pode parecer ininteligível, porém, o uso das tecnologias em favor do gerenciamento de dados e metadados visando a organização e recuperação da informação pode representar uma

alternativa para o ambiente confuso do universo normativo, do ponto de vista dos pesquisadores.

É fundamental, portanto, não apenas a discussão da integração dos instrumentos de gestão de metadados e da representação da informação, mas também o uso das tecnologias na inclusão crítica e didática de instrumentos que favoreçam o rigor às normas e a facilidade de uso pelos pesquisadores, sobretudo em um cenário em que mudanças ocorrem continuamente, decorrentes tanto do desenvolvimento de novas mídias como da complexidade que envolve o tratamento e recuperação da informação.

O uso das tecnologias não pode ser ignorado nesse contexto, sobretudo considerando as funcionalidades da Web Semântica que, em face dos dados ligados, suscita questionamentos sobre a real necessidade do dispêndio de esforços para a normalização de referências, considerando, além da infinidade de estilos bibliográficos distintos, a possibilidade de ampliar os seus índices de encontrabilidade, por meio do tratamento semântico e da interligação de metadados por meio dos dados ligados (*linked data*).

Nesse contexto, entende-se por tratamento adequado à informação aquele em que ao leitor (pesquisador) são oferecidos todos os elementos necessários à identificação da informação representada, de modo a evidenciar as abordagens contidas em um determinado documento, e com as ontologias FaBiO e CiTO há flexibilidade na representação dos metadados. Mais do que uma questão de processo, o atual panorama conduz ao entendimento de que a comunidade científica está alheia aos benefícios, e talvez até à possibilidade, da interoperabilidade aplicada à elaboração de referências.

O mundo da informação se ergueu perante à Ciência da Informação em um cenário em que a integração das tecnologias de gestão de metadados se tornou um requisito para proporcionar ao pesquisador uma experiência positiva no processo de produção, registro, armazenamento, recuperação e uso da informação. Tal cenário sugere que a evolução da organização da informação e do conhecimento passou despercebida por um determinado período pela Ciência da Informação, que agora parece viver um cenário em que a sua reinvenção é inevitável e urgente, como estabeleceu Soergel (2001) em seu trabalho.

REFERÊNCIAS

ALIAGA, F. M.; CORREA, A. D. Tendências em la normalización de nombres de autores em publicaciones científicas. **RELIEVE**, Valencia, v. 17, n. 1, p. 1-10, 2011. Disponível em: <http://www.uv.es/RELIEVE/v17n1/RELIEVEv17n1_0.pdf>. Acesso em: 20 fev. 2018.

BERTHOLINO, M. L. F.; SILVA, V. L. B. da. Normas técnicas de informação e documentação: ABNT versus Vancouver. **Publicatio UEPG: Ciências Biológicas e da Saúde**, Ponta Grossa, v. 14, n. 2, p. 39-44, jun. 2008. Disponível em: <<http://www.revistas2.uepg.br/index.php/biologica/article/view/1275/920>>. Acesso em: 3 jan. 2018.

BERTI JUNIOR, D. W., SOERGEL, D., LIMA, G.A., SANTOS MACULAN, B.C.M. Semiautomatização de relações em tesouros: uma proposta para refinamento de relacionamentos semânticos a partir do Tesouro AGROVOC. **Informação & Informação**, Londrina, v. 22, n. 3, p. 377-404, 2017.

BIZER, C.; HEATH, T.; BERNERS-LEE, T. Linked data: the story so far. **International Journal on Semantic Web and Information Systems**, Hershey, v. 5, n. 3, p. 1-22, 2009.

BREITMAN, K. K. **Web semântica: a internet do futuro**. Rio de Janeiro: LTC, 2014.

KÜCK, G. Tim Berners-Lee's semantic web. **South African Journal of Information Management**, Johannesburg, v. 6, n. 1, mar. 2004. Disponível em: <<https://sajim.co.za/index.php/sajim/article/view/297/288>>. Acesso em: 15 mar. 2018.

PERONI, S. Automating Semantic Publishing, **Data Science**, Amsterdam, v. 1, n. 1-2, p. 155-173, 2017. Disponível em: <<https://content.iospress.com/download/data-science/ds012?id=data-science%2Fds012>>. Acesso em: 7 ago. 2018.

PERONI, S. **Example of use of CiTO #3**. 2015. Disponível em: <https://figshare.com/articles/Example_of_use_of_CiTO_3/1512818/1>. Acesso em: 7 ago. 2018.

PERONI, S.; SHOTTON, D. FaBiO and CITO: ontologies for describing bibliographic resources and citations. **Web semantics: science, services and agents on the world wide web**, Amsterdam, v. 17, p. 33-43, dez. 2012.

PERONI, S.; SHOTTON, D. The SPAR ONTOLOGIES. To appear in Proceedings of the 17th International Semantic Web Conference. Disponível em: <<https://w3id.org/spar/article/spar-iswc2018/>>. Acesso em: 7 ago. 2018.

RDA STEERING COMMITTEE. **Implementation of the LRM in RDA**. 2017. Disponível em: <<http://www.rda-rsc.org/ImplementationLRMinRDA>>. Acesso em: 7 ago. 2018.

RODRIGUES, M. E. F.; LIMA, M. H. T. de; GARCIA, M. J. de O. A normalização no contexto da comunicação científica. **Perspectivas em CI**, Belo Horizonte, v. 3, n. 2, p. 147-156, jul./dez. 1998. Disponível em: <<http://portaldeperiodicos.eci.ufmg.br/index.php/pci/article/view/603/372>>. Acesso em: 15 mar. 2018.

SANTOS, E. A.; MUCHERONI, M. L. VIAF and Open Citations: the cooperative work as a strategy for organization of information at the linked data era. In: INTERNATIONAL ISKO CONFERENCE, 15., 2018, Porto. **Anais eletrônicos...** Porto, 2018.

SOERGEL, D. **The representation of Knowledge Organization Structure (KOS) data.** A multiplicity of standards. Roanoke: JCDL, 2001.

SHOTTON, D. CITO, the Citation Typing Ontology. **Journal of Biomedical Semantics**, London, v. 1, jun. 2010a. Disponível em:
<<https://jbiomedsem.biomedcentral.com/articles/10.1186/2041-1480-1-S1-S6>>. Acesso em: 25 jun. 2018.

SHOTTON, D. **Introducing the Semantic Publishing and Referencing (SPAR) Ontologies.** 2010b. Disponível em:
<<https://semanticpublishing.wordpress.com/2010/10/14/introducing-the-semantic-publishing-and-referencing-spar-ontologies/>>. Acesso em: 23 jul. 2018.