



REPRESENTAÇÃO DOCUMENTAL E DIGITALIDADE

percepções relacionais
entre tradição e inovação

Abecin
EDITORA

**REPRESENTAÇÃO DOCUMENTAL E
DIGITALIDADE**
**percepções relacionais entre tradição e
inovação**

Paula Regina Dal'Evedove
Zaira Regina Zafalon
(organizadoras)

**REPRESENTAÇÃO DOCUMENTAL E
DIGITALIDADE**
**percepções relacionais entre tradição e
inovação**

São Paulo
Abecin Editora
2024

©2024 by Paula Regina Dal'Evedove e Zaira Regina Zafalon (organizadoras)
Direitos desta edição reservados à ABECIN Editora

Rua Maracaju, nº 58, Vila Mariana, São Paulo/SP, CEP 04013-020

ESSA OBRA É LICENCIADA POR UMA

LICENÇA CREATIVE COMMONS



Atribuição – Uso Não Comercial – Compartilhamento pela mesma licença 3.0

É permitido copiar, distribuir, exibir, executar a obra e criar obras derivadas desde que sem fins comerciais e que seja dado o crédito apropriado aos autores e compartilhada sob a mesma licença do original.

Dados Internacionais de Catalogação-na-Publicação (CIP)

R425 Representação documental e digitalidade: percepções relacionais entre tradição e inovação / Paula Regina Dal'Evedove e Zaira Regina Zafalon (organizadoras). – São Paulo: Abecin Editora, 2024.
259 p.

e-ISBN: 978-65-86228-13-7.

Disponível em: <https://portal.abecin.org.br/editora>.

1. Representação documental. 2. Tratamento da informação. 3. Ambientes informacionais digitais. I. Dal'Evedove, Paula Regina, org. II. Zafalon, Zaira Regina, org.

CDU: 02(05)

CDD: 020

Ficha catalográfica: Cleide Oliveira da Silva – CRB-8 8157/SP.

COMISSÃO EDITORIAL E CIENTÍFICA

Editor-chefe: Zaira Regina Zafalon (UFSCar)

Aldinar Martins Bottentuit (UFMA)	José Antonio Frías (USAL, Espanha)
Alessandra dos Santos Araújo (UFS)	José Antonio Moreira González (UC3M, Espanha)
Andréa Pereira dos Santos (UFG)	Manuela Moro Cabero (USAL, Espanha)
Aurora Cuevas-Cerveró (UCM, Espanha)	Márcia Ivo Braz (UFPE)
Célia Regina Simonetti Barbalho (UFAM)	Márcio Bezerra da Silva (UNB)
Danielly Oliveira Inomata (UFAM)	Marta Lúgia Pomim Valentim (UNESP)
Dunia Llanes Padrón (UH, Cuba)	Martha Suzana Cabral Nunes (UFS)
Franciele Marques Redigolo (UFPA)	Meri Nadia Marques Gerlin (UFES)
Helen Beatriz Frota Rozados (UFRGS)	Naira Christofolletti Silveira (UNIRIO)
Henriette Ferreira Gomes (UFBA)	Paulina Szafran (UDELAR, Uruguai)
Ieda Pelógia Martins Damian (USP)	Samile Andréa de Souza Vanz (UFRGS)
Isidoro Gil Leiva (UM, Espanha)	Valéria Martin Valls (FESP/SP)
Ivana Lins (UFBA)	

Revisão: Paula Regina Dal'Evedove e Zaira Regina Zafalon
Diagramação e Editoração: Zaira Regina Zafalon

Normalização: Autores
Capa: Paula Regina Dal'Evedove

O conteúdo dos capítulos e seus dados em sua forma, correção e confiabilidade são de responsabilidade exclusiva dos autores e não representam necessariamente a posição oficial da Editora Abecin. Os originais foram previamente submetidos à avaliação cega pelos pares, membros da Comissão Editorial e Científica desta Editora, tendo sido aprovados para a publicação.

SUMÁRIO

PREFÁCIO	6
<i>Ana Lúcia Terra</i>	
APRESENTAÇÃO	13
<i>Paula Regina Dal'Evedove e Zaira Regina Zafalon</i>	
CAPÍTULO 1 – TECNOLOGIAS DA WEB SEMÂNTICA PARA O ENRIQUECIMENTO DE REGISTROS E COMPLEMENTAÇÃO DE ACERVOS	15
<i>Elisabete Gonçalves de Souza e Darlene Alves Bezerra</i>	
CAPÍTULO 2 – METADADOS PARA REPRESENTAÇÃO DE CONJUNTOS DE DADOS DE PESQUISA E OS PRINCÍPIOS FAIR	54
<i>Caterina Marta Groposo Pavão e Cleusa Pavan</i>	
CAPÍTULO 3 – CATALOGAÇÃO DE ASSUNTO: REFLEXÕES DE TERMINOLOGIA DESDE A DÉCADA DE 1960	87
<i>Rosali Fernandez de Souza e Naira Christofolletti Silveira</i>	

CAPÍTULO 4 – REPRESENTAÇÃO DA INFORMAÇÃO DE OBRAS FOTOGRÁFICAS EM AMBIENTE DIGITAL: MUSEU DE ARTE DE SÃO PAULO ASSIS CHATEAUBRIAND (MASP)	108
<i>Patricia Andreia Lira e Giovana Deliberali Maimone</i>	
CAPÍTULO 5 – TEORIA CRÍTICA: UM OLHAR PARA A ORGANIZAÇÃO DO CONHECIMENTO NA WEB	126
<i>Luciane Paula Vital</i>	
CAPÍTULO 6 – HIBRIDIZAÇÃO DAS FOLKSONOMIAS E INSTRUMENTOS DE CONTROLE TERMINOLÓGICO PARA A REPRESENTAÇÃO DO CONHECIMENTO NA WEB: TENDÊNCIAS E DESAFIOS	141
<i>Raimunda Fernanda dos Santos</i>	
CAPÍTULO 7 – RESGATE DOS FUNDAMENTOS DA CATALOGAÇÃO DE ASSUNTO: PRINCÍPIOS, ENTRADAS PADRONIZADAS E REMISSIVAS EM FOCO NA ANÁLISE DO SITE DA AMAZON	166
<i>Brisa Pozzi de Sousa</i>	
CAPÍTULO 8 – A INSTÂNCIA DE ASSUNTO NO CONTEXTO DIGITAL: EM CENA O TRATAMENTO TEMÁTICO DA INFORMAÇÃO	189
<i>Lais Pereira de Oliveira</i>	

CAPÍTULO 9 – A PRÁTICA DE INDEXAÇÃO EM CORRESPONDÊNCIAS: PROPOSTA DE MODELO DE LEITURA E REPRESENTAÇÃO DOCUMENTÁRIA	211
<i>Amanda Freitas dos Santos e Márcia Ivo Braz</i>	
CAPÍTULO 10 – UMA CONTRIBUIÇÃO PARA REPRESENTAÇÃO DESCRITIVA DE ARTEFATOS TRIDIMENSIONAIS DA CULTURA POPULAR	230
<i>Karin Herculano Picado e Edilene Toscano Galdino dos Santos</i>	
SOBRE AS AUTORAS E ORGANIZADORAS	255

PREFÁCIO

O livro *“Representação Documental e Digitalidade: percepções relacionais entre tradição e inovação”* aborda a organização e a representação da informação na era digital suscitando estimulantes reflexões para pensarmos a reconfiguração de uma área de pensamento e de atuação tradicional de bibliotecários e outros profissionais da informação.

Num mundo onde se diz que existem mais *bits* de informação do que estrelas no universo, a organização e a representação da informação têm de assumir uma necessária centralidade de modo a garantir a recuperação dos itens informacionais e a preservação de sua memória. A obra dá um valioso contributo para este reposicionamento central da organização e da representação da informação ao abordar um conjunto diversificado e complementar de temas, desde os desafios decorrentes da necessidade de tornar o catálogo da biblioteca interoperável, tirando proveito das tecnologias da Web Semântica, ao explicitar as necessidades de representação dos dados de pesquisa, em repositórios de dados, para o cumprimento dos princípios FAIR (*Findable / Localizável, Accessible / Acessível, Interoperable / Interoperável e Reusable / Reutilizável*), ao problematizar as escolhas relativas à representação temática de fotografias em plataformas digitais ou ao sistematizar propostas para a representação de tipologias documentais específicas, como objetos tridimensionais ou documentos de correspondência. Reflexões sobre conceitos e terminologia, cunhados por

autores clássicos da área da organização da informação, e sobre o aporte de algumas teorias para uma reflexão estruturada da organização da informação, em diferentes contextos, são também colaborações para afirmar a relevância da organização e da representação da informação no ambiente digital.

No Capítulo 1, Elisabete e Darlene desenvolvem a sua reflexão em torno da necessidade de orientar o catálogo para um formato interoperável, compatível com as possibilidades abertas pela Web Semântica e suas tecnologias, como o *Linked Data* e a linguagem/vocabulário RDF. Os catálogos das bibliotecas funcionam ainda como silos de dados sem interligação, inviabilizando pesquisas semanticamente mais completas e ricas que poderiam oferecer novas possibilidades de acesso aos acervos das bibliotecas. Neste capítulo, as autoras oferecem uma abordagem didática que procura aproximar os profissionais das bibliotecas aos estudantes da implementação do *Linked Open Data* como aplicação prática da Web Semântica que amplia as possibilidades de representação dos objetos digitais, permitindo-lhes ultrapassar os limites dos catálogos tradicionais tornando-os interoperáveis e acessíveis. Para exemplificar a sua exposição, as autoras usam um registo do acervo das obras raras da Biblioteca Nacional do Brasil, editando-o com o MarcEdit, de modo a acrescentar campos e agregar links semânticos ao registo MARC21, que convertem depois para o padrão Dublin Core.

No Capítulo 2, Caterina e Cluesa tratam da representação da informação no contexto de repositórios de

dados, evidenciando que esta é uma nova área de trabalho e de pesquisa onde os bibliotecários são atores principais para permitir o acesso aos dados de investigação, cumprindo, assim, um dos princípios da Ciência Aberta. As autoras levam a cabo uma reflexão consistente, baseada na revisão dos conceitos operatórios nucleares de dados de pesquisa, repositórios de dados de pesquisa bem como de metadados e na apresentação sucinta de alguns esquemas de metadados. Este percurso permite-lhes sublinhar, de modo fundamentado, que a representação dos dados de pesquisa, através de metadados, é uma condição preliminar para os tornar localizáveis, acessíveis, interoperáveis e reutilizáveis e, portanto, fazer com que cumpram os princípios FAIR, formulados em 2014.

O Capítulo 3, posiciona-se na análise diacrónica da terminologia da catalogação de assunto cunhada por Eric James Coates, na década de 1960, à luz da sua presença nos Requisitos Funcionais para Dados de Autoridade de Assuntos, de 2010, e na Declaração dos Princípios Internacionais de Catalogação, de 2016. Rosali e Naira concluem que os 32 termos elencados por Eric Coates, na sua obra *Subject catalogues: headings and structure* (1960, 1963, 1969), se apresentam válidos na sua essência conceitual, apesar de a terminologia ter se adaptado ao longo do tempo, muito em função dos ambientes computacionais e do desenvolvimento dos Sistemas de Organização do Conhecimento (SOC).

Patrícia e Giovana discutem, no Capítulo 4, a representação da informação de obras fotográficas em ambiente digital. A partir de uma amostra de fotografias

disponibilizadas no *website* do Museu de Arte Moderna de São Paulo Assis Chateaubriand (MASP), é analisada criticamente a primazia atribuída à autoria das fotografias para fins de organização das coleções e da pesquisa e recuperação dos recursos fotográficos, em detrimento da representação temática. Percebe-se, assim, que o MASP assume que o perfil do seu público é de pessoas experientes em artes, em geral, ou mesmo na própria coleção do museu.

No Capítulo 5, Luciane fundamenta a Teoria Crítica como um ponto de partida e uma lente adequada para refletir, questionar e desenvolver Sistemas de Organização do Conhecimento capazes de desconstruir os espaços de poder da Web e para abrir caminho à participação de sujeitos envolvidos na representação da informação aptos para dar voz aos grupos sociais, às comunidades ou aos territórios tradicionalmente marginalizados ou periféricos e cujas visões são subestimadas na Organização do Conhecimento. Neste texto, somos levados a questionar e a reequacionar o papel dos profissionais da informação na Organização do Conhecimento na Web, pensado como local de encontro e de negociação de perspectivas diversas.

A combinação de folksonomias com estruturas de controle terminológico para melhorar os instrumentos de organização do conhecimento no contexto digital constitui o foco do Capítulo 6. Raimunda Fernanda apresenta uma reflexão proveitosa e esclarecedora sobre diferentes tipos de garantia (literária, cultural, organizacional, entre outras), os quais subsidiam a validação e a inclusão de consensos semânticos nos Sistemas de Organização do Conhecimento. O

texto desenvolve-se, depois, à volta de exemplos retratados na literatura sobre o processo de hibridização das folksonomias com Sistemas de Organização do Conhecimento, permitindo visualizar novas formas de tornar os conteúdos pesquisáveis para os utilizadores, baseando-se na linguagem desses mesmos utilizadores.

O Capítulo 7 aborda o site de negócio eletrônico Amazon à luz dos fundamentos da catalogação de assunto, em específico no que toca aos princípios, entradas padronizadas e remissivas, tal como propostas por Charles Ami Cutter, em 1876. Neste texto, Brisa evidencia as convergências existentes na representação da informação em bibliotecas com aquela que é necessária e praticada em sites de negócio eletrônico, já que ambos os contextos partilham o objetivo de apresentar/divulgar produtos para que possam ser pesquisáveis e localizáveis.

No Capítulo 8, Laís reflete sobre a análise de assunto, com enfoque para uma sistematização da reflexão produzida na literatura no que respeita ao tratamento temático da informação em paralelo com o impacto do crescimento da informação digital. Neste cenário, onde se verifica uma crescente disponibilização de conteúdos digitais, sobressaem os problemas decorrentes da ausência da descrição padronizada desses recursos. A autora conclui aludindo às potencialidades da Web Semântica para criar e partilhar metadados descritivos e temáticos estruturados que potenciem a representação dos conteúdos informacionais atendendo às necessidades dos utilizadores.

No Capítulo 9, Amanda e Márcia propõem um modelo de representação documentária para a indexação de documentos de correspondência, que aplicaram a documentos do acervo virtual da Fundação Joaquim Nabuco. A revisão dos conceitos de representação da informação, indexação e da tipologia documental correspondência permitiu-lhes sistematizar cinco parâmetros para um modelo de indexação de correspondência e desenhar um modelo de leitura e representação baseado em cinco questões (Quem escreve? Quem lê? O que lê? Onde? Quando?).

A obra termina com o Capítulo 10 que aborda a representação descritiva de artefatos tridimensionais da cultura popular. Reconfigurando propostas anteriores e a correspondência e/ou equivalência com o MARC21 e o AACR2, é proposto um modelo de descrição com 21 campos e apresentados os resultados da sua aplicação numa amostra de peças selecionadas. Karin e Edilene abordam estes objetos tridimensionais como documentos e evidenciam a importância da sua representação descritiva para a disponibilização num sistema de recuperação da informação.

A obra será útil para todos aqueles que procuram reflexões teóricas sobre diferentes perspectivas da representação da informação, mas também para outros, como profissionais e estudantes de Ciência da Informação que procuram relatos de experiências inovadoras ou em cenários pouco habituais para a representação da informação. Com a leitura dos diferentes textos, nos é oferecido um interessante panorama de teorias, princípios, normas e ferramentas atualmente direcionadas para a organização da informação, no

contexto mais tradicional das bibliotecas, mas igualmente noutros ambientes, permitindo uma viagem entre a tradição e a inovação numa das áreas nucleares da Ciência da Informação.

Ana Lúcia Terra

Universidade de Coimbra

APRESENTAÇÃO

A presente coletânea reúne pesquisadoras brasileiras que investigam e participam ativamente na construção de conhecimentos e reflexões que envolvem a organização e a representação da informação na era digital. As autoras convidadas foram encorajadas a refletir, reavaliar e debater a relação dimensional entre Representação Documental e ambiente *Web*.

A ideia de produzir esta obra decorre do atravessamento da digitalidade na Representação Documental. Os primeiros esforços das organizadoras em agregar conhecimentos, conceitos e aplicações da relação entre Tecnologia, Informação e Representação ocorreu em 2017, ano em que foi realizado o I EnReDo - Encontro de Representação Documental, evento científico promovido pelo Programa de Pós-Graduação em Ciência da Informação da Universidade Federal de São Carlos (PPGCI/UFSCar) que incitou debates acerca da consolidação de discussões e ações que envolvem a organização e a representação dos recursos informacionais, bem como os processos de gerenciamento de dados em meio automatizado e seu uso.

A dinâmica evolutiva das técnicas e teorias que envolvem a organização e a representação da informação no ambiente digital requer uma variedade e multiplicidade de visões e perspectivas que introduzam debates atuais e inovadores. Debates sobre a Representação Documental, no correr da história, das ideias, estão em constante

desenvolvimento, e são cada vez mais relevantes para se compreender as manifestações da informação na *Web*.

A Representação Documental na digitalidade nos lança na interrogação, requer novos pensares, suscita novos olhares. Assim sendo, a articulação entre práticas tradicionais e inovadoras para o tratamento da informação são exigíveis no ambiente digital. Considerando-se os novos tempos, novas demandas, propostas concretas e reflexões instigantes são apresentadas por estudiosas do campo da informação para contribuir e responder às demandas dessa integração. Oportunizam, portanto, compreensão sobre o impacto, potencialidades e desafios do ambiente digital para a organização e a representação da informação.

Acreditamos que esta obra oferece um olhar transversal sobre a Representação Documental, especialmente em relação às questões que influenciam e alteram, atualmente, as formas de tratamento da informação no ambiente digital. Compreendendo que esses debates podem ser vistos e analisados a partir de diferentes perspectivas, convidamos a comunidade científica e profissional da Ciência da Informação à reflexão e à ampliação das contribuições aqui compartilhadas.

Agradecemos às autoras que contribuíram com discussões que enriquecem a compreensão das dinâmicas relativas à Representação Documental no ambiente digital, plural e diverso.

As organizadoras

CAPÍTULO 1

TECNOLOGIAS DA WEB SEMÂNTICA PARA O ENRIQUECIMENTO DE REGISTROS E COMPLEMENTAÇÃO DE ACERVOS

Elisabete Gonçalves de Souza

Darlene Alves Bezerra

1 INTRODUÇÃO

A diversidade de recursos informacionais transformou a *web* em uma grande plataforma de conteúdos, tornando-a o principal lugar para a busca de informações. Ciente dessas transformações a International Federation of Library Associations and Institutions (IFLA) vem, desde 2009, atualizando os Princípios da Catalogação no sentido de aproximar as ferramentas bibliográficas, como os catálogos, dos novos recursos proporcionados pelas tecnologias digitais, com vista a ampliar as possibilidades de acesso aos registros e ao conteúdo dos documentos por parte dos usuários, sendo estes definidos “[...] como qualquer indivíduo que busque o catálogo e utilize seus dados” (IFLA, 2016).

O catálogo deve ser pensado como uma ferramenta de comunicação entre o acervo e os usuários, sendo os seus interesses e necessidades o foco do trabalho do bibliotecário. As tarefas dos usuários – encontrar, selecionar, identificar, obter e navegar, são conceitos fundamentais para a concepção dos catálogos.

O presente capítulo centra-se em discutir as possibilidades de se otimizar as demandas postas pelas tarefas “obter” definida como “[...] acesso a um item descrito por meio de compra, empréstimo etc. ou aceder (acessar) eletronicamente a um item por meio de uma ligação em linha a uma fonte remota” e “navegar”, no catálogo ou para além dele através da “[...] organização lógica da informação bibliográfica e da apresentação das formas (IFLA, 2016), identificando “[...] as relações entre obras, expressões, manifestações e itens”. (Silveira; Tálamo, 2009, p. 113).

A centralidade dos usuários nos estudos da catalogação vem exigindo do bibliotecário o comprometimento não só com a organização dos dados bibliográficos e de autoridade e da apresentação das relações entre as entidades bibliográficas no universo do catálogo, mas sinalizam para a possibilidade de encontrar, identificar, selecionar e obter obras em outras instituições que compartilham as mesmas ou outras expressões e manifestações dessas obras.

Marcondes (2016) em estudo sobre a interoperabilidade entre acervos destaca a importância das tecnologias *web* nos processos de complementaridade dos acervos com vista a oferecer melhores serviços aos seus usuários¹. Trata-se de esforços no sentido de alcançar a

¹ Na gestão dos catálogos a interoperabilidade “[...] está presente na troca de metadados, na sua agregação em uma base de dados comum e na operação desta base por programas gerenciadores” (Marcondes, 2016, p. 62). Tais processos vêm tornando-se mais significativos por conta do uso da linguagem RDF nos processos de representação, permitindo agregar dados aos registros ampliando as possibilidades de navegação para além do catálogo no sentido de abri-lo para receber dados de outras instituições.

interoperabilidade por meio do enlaçamento de dados significativos usando a tecnologia do *Linked Open Data* (LOD) - dados abertos interligados.

Parte-se do pressuposto de que a *web* reúne informações significativas e estruturadas, muitas delas com identificadores persistentes, permitindo-nos completar os registros bibliográficos e oferecer aos usuários mais opções de acesso a conteúdo, além de abrir o catálogo a outras instituições que possuem acervos digitais abertos.

Os objetivos são: mostrar que recursos digitais disponíveis em outras bases de dados podem complementar acervos, além de oferecer aos usuários outras possibilidades de acesso aos mesmos; avaliar estratégias que visem a ampliação de campos do formato *Machine Readable Cataloging* (MARC21) com vista a criar meios para a integração de acervos digitais publicados na *web* usando a tecnologia LOD.

Trata-se de pesquisa exploratória de abordagem descritiva, pautada no levantamento de artigos sobre o tema para fundamentar teórica e metodologicamente o problema e buscar respostas plausíveis que possam ser aplicadas pelos bibliotecários na gestão de seus catálogos proporcionando aos seus usuários outras estratégias de navegação e acesso. Para tanto lança-se mão do programa MarcEdit, ferramenta que realiza importação de metadados MARC21, assim como a edição e acréscimos de novos campos e subcampos, permitindo-nos integrar e/ou completar dados em um registro e reintegrá-lo ao acervo.

Em termos epistemológicos tais questões nos fazem rever os paradigmas da catalogação enquanto uma atividade voltada para a organização de catálogos com vista a atender os objetivos da instituição e a preservação do acervo e não necessariamente os interesses dos usuários. O avanço das tecnologias de informação vem permitindo-nos relativizar a questão custodial em prol de ações que ampliem o direito à informação, sejam elas realizadas através de políticas de digitalização, da criação de bibliotecas digitais e mais recentemente por meio da interoperabilidade entre acervos.

Para demonstrar como as tecnologias *web* podem agregar valor (semântico e cultural) aos catálogos realiza exercício de edição dos metadados de uma obra da Coleção de Obras Raras da Biblioteca Nacional, completando ou substituindo as informações dos campos e/ou subcampos MARC por *links* permanentes referentes a informações sobre autor e à localização da obra. Os dados foram coletados no *Virtual International Authority File* (VIAF) e no acervo da Biblioteca Digital Europeia. Discorre também sobre a possibilidade de se selecionar parte do acervo, enriquecê-lo e convertê-lo para ao padrão *Dublin Core* e depositá-lo em coleções especiais a serem criadas em repositórios institucionais.

Para a discussão dessas questões organizamos o estudo em três seções, as quais abordam os dados abertos interligados e a função dos catálogos, seus registros e os limites impostos pela sintaxe descritiva estruturada pelo padrão MARC21 e a necessidade de torná-lo um formato interoperável, compatível com as possibilidades abertas pela

Web Semântica e suas tecnologias, como o *Linked Data* e a linguagem/vocabulário RDF.

2 DADOS ABERTOS INTERLIGADOS

A Web Semântica trouxe novas perspectivas para a atividade de organização e representação do conhecimento no meio digital, visto que os sistemas de organização do conhecimento têm sido desenvolvidos de uma forma mais sofisticada de modo a incorporar um alto grau de inteligência e, portanto, se mostram mais adequados para interligar dados.

Imaginada como uma extensão da web atual por Berners-Lee, Hendler e Lassila (2001), a Web Semântica possibilita por meio de tecnologias, linguagens e padrões que as máquinas executem de maneira eficiente o trabalho cooperativo com os seres humanos. Contudo, a heterogeneidade semântica dos sistemas de organização e representação do conhecimento é vista como um desafio para o alcance da interoperabilidade.

Segundo Zaidan e Bax (2013, p. 47), o termo interoperabilidade é definido pelo *World Wide Web Consortium* (W3C) como “[...] a capacidade de dois ou mais sistemas de interagir e trocar dados e informações, de acordo com métodos definidos e objetivando resultados esperados”. O W3C, consórcio internacional fundado por Tim Berners-Lee, tem se dedicado a desenvolver padrões que objetivam garantir o crescimento da *web* e sua melhoria em termos de interoperabilidade, oferecendo maior expressividade para os dados.

Nesse sentido, visando a melhoria da Web Semântica, padrões, linguagens, esquemas, modelos de dados, ontologias e modelos conceituais têm sido desenvolvidos com o objetivo de promover maior conectividade entre os dados. O consórcio W3C tem liderado essa iniciativa e recomendado o uso de padrões certificados que contribuam para tornar a *web* uma plataforma de conectividade a serviço da humanidade.

Para isso, parte-se do princípio de que todo recurso *web* necessita de um *Uniform Resource Identifier* (URI), uma cadeia de caracteres compacta que nomeia univocamente todo e qualquer recurso (conteúdo). Um URI funciona como um endereço único de um conteúdo publicado na *web*, ele é permanente, o que possibilita encontrá-lo sempre no mesmo endereço.

Desde 1999 a linguagem declarativa *Resource Description Framework* (RDF) é o padrão recomendado pelo W3C. A base de escrita da RDF é a *Extensible Markup Language* (XML) e a representação de um recurso é feita por meio de um conjunto de triplas, cuja estrutura sintática consiste na sentença: *sujeito, predicado e objeto*.

Heath e Bizer (2011) esclarecem que o *sujeito* de uma tripla é um URI que identifica um recurso. Já o *objeto* pode ser um valor literal simples, como uma *string* (cadeia de caracteres), um número, uma data ou, ainda, o URI de outro recurso que está de alguma forma relacionado ao *sujeito*. Por fim o *predicado*, elemento que indica que tipo de relação existe entre o *sujeito* e o *objeto*, sendo também identificado por um URI ou por uma *string*.

Devido a RDF apresentar uma semântica simplificada, padrões mais expressivos foram desenvolvidos de modo a fornecerem uma representação mais abrangente, como o *Resource Description Framework Schema* (RDFS), vocabulário que modela os dados RDF. O RDFS é uma extensão do vocabulário RDF básico, o qual possibilita a definição de classes e tipos em RDF. Outros vocabulários também foram desenvolvidos: o *Ontology Web Language* (OWL) e o *Simple Knowledge Organization System* (SKOS).

O OWL possibilita a descrição de aspectos semânticos dos termos utilizados e seus relacionamentos, o que favorece uma representação mais abrangente do RDF e melhoria das questões de interoperabilidade. Já o SKOS é um vocabulário para expressar hierarquias conceituais, muitas vezes referidas como taxonomias. “[...] Coletivamente, RDFS, OWL e SKOS fornecem um *continuum* de expressividade” (Heath; Bizer, 2011, p. 57). Esses padrões reunidos permitem que se alcance a desejada Web Semântica, onde tecnologias semânticas facilitem a publicação, o consumo, o reuso e a ligação de dados na *web*, promovendo interoperabilidade entre sistemas interconectados.

O termo “tecnologias semânticas” representa “[...] um conjunto de famílias que buscam o significado dos dados e da informação” (Zaidan; Bax, 2013, p. 44). Isso caracteriza uma *web* mais sofisticada que privilegie a consistência, o significado e a integridade dos dados. Nela, a ênfase não se encontra em uma página apontando para outra, mas em um dado

apontando para outro dado, usando referências globais, como o *Linked Data*².

Segundo Heath e Bizer (2011, p. 7, grifo nosso), “[...] o termo *Linked Data* refere-se a um conjunto de *boas práticas* para publicar e interligar dados estruturados na *web*”, de modo a promover a percepção de valores de dados derivados. Foram introduzidas por Berners-Lee na arquitetura *web* em 2006. Posteriormente, tornaram-se conhecidas e adquiriram o *status* de princípios, os quais são: usar URI como nomes para recursos; usar URI e HTTP³ para que as pessoas possam encontrar esses nomes; ou, quando alguém procurar por um URI, devem ser fornecidas informações úteis, usando os padrões RDF e SPARQL⁴ e, por fim, incluir *links* para outros URI de forma que outros recursos possam ser descobertos.

Esses quatro princípios foram denominados *Linked Data Principles* e fornecem diretrizes “[...] para publicar e interligar dados usando a infraestrutura da *web* aderindo à sua

² Em 2006, Tim Berners-Lee cunhou o termo *Linked Data* para se referir a capacidade de uma *web* conectar dados relacionados, tornando-os mais úteis de modo a contribuir para a diminuição de barreiras para ligá-los. No ano seguinte (2007), o W3C iniciou o projeto aberto comunitário mundial denominado *Linked Open Data* (LOD) que visa à publicação de vários conjuntos de dados (*datasets*) de forma que as ligações sejam possíveis entre eles. O modelo de dados LOD “[...] fará parte da infraestrutura básica da *web* disponibilizando de forma acessível os dados interligados e acomodando mais facilmente as aplicações que ainda mantêm os dados internamente” (Zaidan; Bax, 2013, p. 45).

³ Segundo Tammaro e Salateli (2008), o *Hypertext Transfer Protocol* (HTTP) é um protocolo desenvolvido por Tim Berners-Lee na década de 1990 como estrutura de sustentação de todo o universo *web*.

⁴ *Simple Protocol and RDF Query Language* (SPARQL) é uma linguagem que permite consultar triplas RDF (*Free tools, information and resource for the Semantic Web*) (Yankulov, 2024).

arquitetura e padrões” (Bizer; Heath; Berners-Lee, 2009). As diretrizes determinam que os dados em RDF devem ser: (i) abertos (e não proprietários), pois devem ser acessados por meio de uma ilimitada variedade de aplicações; (ii) modulares, uma vez que não necessitam de planejamento prévio para combinar com outros dados; e (iii) escaláveis, pois uma vez que já existem dados em RDF, a adição de novos dados tende a ser facilitada (Zaidan; Bax, 2013).

A construção de sistemas de organização e representação do conhecimento compatíveis com as tecnologias LOD permitirá a complementação de acervos por meio da integração com outras fontes, como a Wikipedia, DBpedia, Geonames, Europeana etc. Bases de dados que possuem licença aberta e fornecem descrições em RDF de itens, com *links* de entrada e saída de recursos. O controle semântico dos recursos informacionais, através de vocabulários controlados e outras linguagens documentárias, é um dos desafios dos estudiosos da área de organização da informação e do conhecimento. Estes também devem buscar compatibilizá-los com as tecnologias *Linked Data* a fim de expandir o alcance aos recursos alocados em acervos, visto que o usuário e suas necessidades deve ser o principal alvo no desenvolvimento de sistemas de organização e representação do conhecimento.

As relações semânticas são fundamentais para agregar sentido aos recursos, ajudando-nos a identificar o criador de uma obra, sua natureza, quem a produziu, seu conteúdo etc. Em ambientes digitais, as relações semânticas expressam o grau de capacidade dos agentes de *softwares* realizarem

“inferências” automáticas “[...] descobrindo conhecimento não diretamente explicitado, com base nessas relações, seguindo os ‘links’ entre os nós de uma rede semântica” (Marcondes, 2016, p. 65).

Com a tecnologia do LOD a semântica dos conteúdos publicados na *web* tem que ser agregada a dispositivos como modelos conceituais, ontologias e/ou vocabulários semânticos e representados “[...] em linguagem legível por programas e associados às triplas RDF, que especificam univocamente o significado dos predicados e objetos de cada tripla” (Marcondes, 2016, p. 65).

3 OS CATÁLOGOS NO CONTEXTO DA WEB

O catálogo, conforme destacou Cutter, deve ser um instrumento que permita que o usuário encontre um recurso por meio de seu autor, título e/ou assunto; mostrar todos os recursos que a biblioteca possui e orientá-lo na escolha, embasado pela sua edição, tipologia e conteúdo (Fiúza, 1987; Mey; Silveira, 2009).

Corroborando esta definição o entendimento do catálogo como um conjunto de representações de diferentes itens documentais; um recurso cognitivo que permite a um usuário, por meio de inferências, “[...] decidir acerca da validade de obter e realizar a leitura do documento completo, [o que implica] em uma significativa economia de energia em comparação com a obtenção e leitura do documento original” (Marcondes, 2001, p. 67).

Com a disponibilização do *Online Public Access Catalog* (OPAC) o alcance dos catálogos ampliou-se, possibilitando tanto a importação como a exportação de registros entre os sistemas de bibliotecas otimizando os processos de busca, recuperação, assim como o uso e o reuso da informação. No que diz respeito à descrição seu “[...] cerne deslocou-se do item para o usuário, visando permitir-lhe as tarefas de encontrar, identificar, selecionar e obter uma ‘entidade’ adequada aos seus propósitos” (Mey, 2003, p. 1).

Mas no ambiente *web* as tarefas de busca por informações são realizadas de forma automática pelos agentes de *softwares*, o que exige que as formas de representação sejam compreensíveis tanto por humanos como por computadores. Isso vem implicando no desenvolvimento e no aprimoramento de uma nova sintaxe para a estruturação dos metadados que seja semanticamente compatível com as tecnologias digitais disponíveis. Isso nos coloca diante de um novo entendimento do que seja um catálogo, compreendendo-o como uma ferramenta que não se restringe a descrever recursos presentes em um acervo, podendo alcançar e difundir acervos de outras instituições que estão em acesso aberto na *web*. O catálogo torna-se, de fato, um poderoso meio de comunicação a veicular mensagens “[...] sobre os registros do conhecimento, de um ou vários acervos, reais ou ciberespaciais, apresentando-as com sintaxe e semântica próprias e reunindo os registros do conhecimento por semelhanças [...]” (Mey; Silveira, 2009, p. 12), não apenas para os usuários do acervo, mas para todos que o consulte, explicitando suas entidades, atributos e relacionamentos.

Para o campo da descrição se faz necessário uma padronização de linguagens e de metadados descritivos e de regras comuns “[...] sobre como armazenar dados e descrever a informação armazenada e que esta possa ser ‘consumida’ por outros usuários humanos ou não, de maneira automática e não ambígua” (Souza; Alvarenga, 2004, p. 134).

Estruturada em campos e subcampos a informação bibliográfica de um registro MARC tem sua semântica assegurada pela ordem dos elementos descritivos e por sua respectiva pontuação, mas fora desse ambiente (como na *web*) essa descrição torna-se ambígua. Mey e Silveira (2009, p. 5) explicando a função da *International Standard Bibliographic Description (ISBD)* na narrativa descritiva dizem: “[...] A sintaxe se constitui de posição e pontuação. A semântica dá significado aos termos em sua posição e pela pontuação precedente”. Por exemplo:

[...] o número ‘1984’ pode representar: uma data de publicação, o número de páginas de um livro, o título de uma obra ou o pseudônimo de um autor, entre outras alternativas; a diferença entre as diversas informações consiste na posição em que cada uma delas se encontra e na pontuação que as precede, ou seja, na sintaxe, que expressa a semântica: 1984 / George Orwell ... [título de uma obra] ... Rio de Janeiro : Record, 1984 [data de publicação] ... 1984 p. [número de páginas] XYZ / 1984 ... [pseudônimo de um autor] (Mey; Silveira, 2009, p. 5).

A representação dos recursos em um catálogo envolve decisões embasadas em códigos e normas de catalogação como o *Anglo-American Cataloguing Rules (AACR)*,

International Standard Bibliographic Description (ISBD) e *Resource Description Access* (RDA), que implicam na forma e na ordem de descrição dos dados bibliográficos e de autoridade, contribuindo para sua identificação pelos usuários, pessoas e sistemas e, dependendo dos formatos e protocolos adotados os dados poderão ser intercambiados. Nota-se, portanto, que o trabalho de descrição não é limitado a elementos sintáticos, mas envolve também elementos semânticos, o que nos permite ampliar as opções de serviços a serem oferecidos aos usuários. Nesse processo, a linguagem representacional exige uma estrutura sintática que permita organizar “[...] os metadados de forma que eles sejam confiáveis e legíveis por máquinas” (Serra; Santarém Segundo, 2017, p. 170), de modo que possamos abstrair a semântica expressa no registro documentário. Ou seja: compreender o significado dos dados, permitindo que as informações sejam identificadas e interpretadas.

No formato MARC, por exemplo, a sintaxe é definida pelos campos, subcampos e indicadores, elementos que norteiam a estruturação dos registros, sua forma discursiva legível por máquina. Por meio desse arranjo, a semântica aflora; contextualiza sobre o que o dado discorre, identificando os valores dentro dos campos.

Zafalon e Santos (2012, p. 14) explicam que a sintaxe se relaciona à forma, à disposição em que cada elemento deva ser descrito; determina relações formais entre a representação de cada uma das partes do documento representado, que reunidas nos permite identificar o recurso e conhecer sua estrutura semântica. De acordo com as autoras “[...] a

semântica de registros bibliográficos relaciona-se tanto ao significado estabelecido entre cada um dos elementos representados [...] quanto o significado entre o elemento definidor e o valor dado para expressar este elemento [...]”. Significa que:

[...] o registro bibliográfico de um documento só pode assumir o significado de representação daquele documento dado o sentido que pode ser estabelecido entre representado e representante, ou, entre obra e manifestação; pelo significado entre o elemento a ser representado e o seu valor (dado, por sua vez, pela sintaxe), e uma metasemântica, no sentido das relações de metalinguagem que podem ser estabelecidas entre diferentes registros (Zafalon; Santos, 2012, p. 14).

A arquitetura dos programas gerenciadores de catálogos tem papel preponderante na semântica dos conteúdos. A distribuição dos dados em campos e subcampos de forma padronizada, seguindo uma sintaxe previamente estabelecida por meio de normas e padrões, faz com que uma cadeia de caracteres ganhe sentido para o usuário permitindo-lhe encontrar o título, o autor ou o assunto de um recurso. No entanto, como destaca Marcondes (2016), “[...] ao serem publicados diretamente na *web* através das tecnologias de dados abertos interligados estes conteúdos tornam-se independentes dos programas gerenciadores de catálogos”.

Nos catálogos estruturados em MARC a descrição é feita de forma textual e linear, permitindo a leitura por pessoas, mas não por agentes de *softwares* tornando os registros não interoperáveis com os novos padrões da Web

Semântica, mas isso não inviabiliza possibilidades de enriquecimento desses registros agregando-os a *links* semânticos⁵ e através deles integrar conteúdos digitais de acervos heterogêneos custodiados por Bibliotecas, Arquivos e Museus. Segundo Marcondes (2016, p. 65-66) “[...] isto permitirá a estas instituições tirarem partido das sinergias e complementaridades entre seus acervos e oferecerem melhores serviços – prover informação cultural curada, tratada, contextualizada, de alta qualidade, aos seus usuários. O autor termina suas reflexões destacando que a Web Semântica é “[...] uma evolução da Web atual (sintática) em direção a mais amplas possibilidades de interação e de recuperação semântica da informação” (Marcondes, 2012, p. 176), que ambiciona interligar os dados abertos criando ecossistemas que permitam que os agentes (humanos e máquinas) possam realizar tarefas e descobrir conteúdo.

A linguagem RDF é o padrão básico da proposta de LOD. Nas assertivas em RDF tanto o recurso (o sujeito da assertiva) quanto o predicado e o objeto podem ser identificados univocamente em todo o espaço da *web* por *links* diferentes dos convencionais, *links permanentes* que expressem a natureza ou “semântica” da ligação entre os recursos – os URI (Marcondes, 2012). Significa que, ao invés de termos *links* para páginas hipertextuais teremos *links* para os recursos.

⁵ “[...] um link semântico seria, por exemplo, ligar o registro do arquivo digital de um filme numa cinemateca digital com a imagem digital de uma pintura no acervo de um museu, que inspirou o filme” (Marcondes, 2016, p. 66).

A proposta do LOD, como vem destacando Marcondes (2012, 2016), acena com a possibilidade de interligarmos acervos em arquivos, bibliotecas e museus digitais através de tecnologias *web*. O grande desafio é poder usufruir, ainda que parcialmente, dessas tecnologias para enriquecer ou completar com dados significativamente relevantes os acervos interferindo na sua estrutura de metadados MARC, agregando campos que nos permitam conhecer mais sobre os criadores, produtores e custodiadores dos recursos, assim como ter acesso às diferentes expressões e manifestações dos recursos disponíveis em acesso aberto.

3.1 O padrão MARC: em busca da interoperabilidade semântica

Os códigos e normas, como o AACR2r, a ISBD, a RDA fixam as regras para a descrição bibliográfica, mas o intercâmbio dos dados depende de formatos e protocolos. Dentre esses destaca-se o formato MARC.

A Library of Congress (LC) desenvolveu o formato MARC na década de 1960 com o objetivo de facilitar a rotina bibliotecária de desdobramento de fichas catalográficas. Nota-se que o propósito inicial não foi o intercâmbio de registros, “[...] mas sim atender às necessidades descritivas de catálogos em papel, transferindo aos computadores as informações contidas nas fichas analógicas, sem, contudo, explorar recursos tecnológicos” (Serra; Santarém Segundo, 2017, p. 172). Mas essa situação já não se aplica ao cenário atual e vem exigindo a busca de outros padrões que suporte a estrutura MARC, mas agreguem as novidades proporcionadas pelas ferramentas *web*. Essa questão colocou em cena a discussão

em torno de dois padrões de metadados, o *MarcOnt Initiative* e o *Bibliographic Framework Initiative* (BIBFRAME).

No entanto, como nos lembra Castro e Santos (2008), o desenvolvimento de padrões semânticos “[...] não pode negligenciar padrões e formatos existentes ou nunca serão amplamente reconhecidos”. No caso específico das bibliotecas, o formato MARC21 é ainda o principal padrão utilizado pela maioria das instituições, por isso “[...] A ontologia que aspira ser seu sucessor (ou ao menos contrapartes semânticas) tem que oferecer não menos capacidade na descrição” (Synak, 2005 *apud* Castro; Santos, 2008, p. 4).

A interoperabilidade dos sistemas de informação é o elemento-chave para a integração dos acervos. Mas a complexidade da estrutura MARC21 é um desafio. Essa questão é abordada por Castro e Santos (2008) em artigo sobre o padrão de descrição semântica MarcOnt, cuja proposta é proporcionar uma descrição semântica que contemple a profundidade da descrição suportada pelo MARC, que há mais de 50 anos vem atendendo os propósitos da Biblioteconomia. Os autores consideram que o padrão “[...] pode possibilitar uma melhora na explicitação das informações de forma a permitir uma compreensão pela máquina nas solicitações de busca para o usuário final” (Castro; Santos, 2008, p. 4), inferência possível porque o padrão utiliza ferramentas tecnológicas subjacentes à Web Semântica, como a linguagem RDF para o armazenamento das suas descrições o que torna seus recursos compartilháveis e interoperáveis com outras bibliotecas digitais e sistemas de informação que fazem uso de padrões de metadados heterogêneos.

Em estudo que teve como campo empírico a Biblioteca Digital Semântica JeromeDL, estruturada no padrão MarcOnt, os autores observaram que o sistema não absorve a complexidade da descrição bibliográfica concentrada em um registro MARC21, o que fez com que, até o momento em que o estudo foi realizado, seus gestores utilizassem apenas os campos de autoria (1XX) e título (245) para a busca e a recuperação dos recursos informacionais. Esses limites precisam ser superados, mas pelo que parece ainda permanecem.

Para que o padrão de descrição semântica MarcOnt, se torne pleno é fundamental que absorva a égide dos preceitos e princípios que orientam a Catalogação Descritiva, pois só assim será possível melhorar sua performance garantindo de forma plena “[...] a interoperabilidade semântica entre os padrões de metadados vislumbrados em JeromeDL e outros sistemas e ambientes informacionais digitais, de forma a garantir a apresentação dos recursos para a preservação, o uso e o reuso das informações para o usuário final” (Castro; Santos, 2008, p. 4).

A busca pela interoperabilidade semântica levou a LC a lançar, em 2012, o BIBFRAME, um modelo de dados que fornece uma base para descrição bibliográfica na *web*, sendo para isso fortemente estruturado em técnicas *Linked Data*. Em termos conceituais o BIBFRAME é uma evolução do formato MARC21. Sua principal diferença em relação ao formato anterior é o foco na Web Semântica, em relacionar dados bibliográficos por meio de URI e uso do vocabulário RDF como modelo básico de descrição de dados.

Projetado para integrar e envolver a comunidade de informação e, mais especificamente atender as necessidades das bibliotecas, a iniciativa BIBFRAME busca novas formas de diferenciar o conteúdo conceitual e suas manifestações físicas/digitais, identificar inequivocamente as entidades de informação (por exemplo, agentes) e alavancar e expor as relações entre as entidades (Library of Congress, 2021).

A versão original do BIBFRAME – BIBFRAME 1.0 – era composta de quatro classes/entidades principais: *work*, *instance*, *authority* e *annotation*. Atualizado em abril de 2016, o modelo passou a ser denominado BIBFRAME 2.0, o qual organiza a informação descritiva em três classes/entidades de alto nível: *work* [obra], *instance* [instância] e *item* [item]. Uma *work* BIBFRAME é o mais alto nível de abstração, a qual identifica a essência conceitual de algo; já uma *instance* BIBFRAME reflete a incorporação material de uma *work*; enquanto um *item* BIBFRAME é uma cópia real (física ou digital) de uma *instance*.

Os conceitos-chave que se relacionam com as classes/entidades principais são: *agents* [agentes], *subjects* [assuntos] e *events* [eventos]. Os *agents* são pessoas, organizações, jurisdições etc. associados a uma *work* ou *instance* por meio de funções como autor, editor, artista, fotógrafo, compositor, ilustrador etc. Uma *work* pode ser “sobre” um ou mais conceitos, sendo cada um deles considerado um “*subject*”. Os conceitos podem ser assuntos que incluem tópicos, lugares, expressões temporais, eventos, obras, instâncias, itens, agentes etc. Os *events* são ocorrências,

cujo registro pode ser o conteúdo de uma *work*. (Library of Congress, 2021).

O vocabulário BIBFRAME é a chave para a descrição dos recursos. Possui um conjunto de classes e propriedades RDF que incluem as três principais listadas acima, bem como várias classes adicionais, muitas das quais são subclasses das principais. As propriedades descrevem as características dos recursos e os relacionamentos entre eles (Library of Congress, 2021).

Comparando-os, pode-se dizer que uma classe BIBFRAME identifica um tipo de recurso - da mesma forma que um campo MARC pode agrupar um único conceito -, já as propriedades servem como um meio para descrever melhor um recurso BIBFRAME - assim como os subcampos MARC identificam mais especificamente os aspectos dos elementos descritos.

Diferente do MarcOnt o novo modelo suporta a complexidade dos registros MARC além de associá-lo às ferramentas web, provocando os relacionamentos entre recursos, tais como: *work-to-work* [obra-para-obra], *work-to-instance* [obra-para-instância], *work-to-agent* [obra-para-agente], que ocorrem por meio de identificadores controlados e permanentes - um URI - para coisas, isto é, pessoas, lugares, idiomas etc. (Library of Congress, 2021).

É interessante observar que algumas dessas ideias já são empregadas no formato MARC por meio de códigos geográficos e de linguagem. Mas no BIBFRAME esses aspectos são considerados uma norma e não uma exceção.

Em síntese: pode-se considerar o modelo a porta de entrada da comunidade de bibliotecas na *web* de dados, onde os *links* entre coisas são fundamentais, como preconiza o *Linked Data*. O modelo procura equilibrar as necessidades dos que registram uma descrição bibliográfica detalhada, ou dos que descrevem outros materiais culturais, ou, ainda, dos que não exigem um nível de descrição detalhado superando, assim, as expectativas depositadas no MarcOnt.

O BIBFRAME é uma nova maneira de apresentar e trocar dados bibliográficos de forma consistente, pois preza por todos os aspectos (teóricos e metodológicos) da descrição bibliográfica acomodando diferentes modelos de conteúdo e regras de catalogação, além de explorar novos métodos de entrada de dados e protocolos de intercâmbio.

Apesar de todas essas virtudes, alguns componentes do modelo ainda estão em discussão e em desenvolvimento, ou seja, é um trabalho ainda em andamento. Mas, pelos objetivos traçados pela LC é possível perceber que o modelo possibilitará a comunidade de bibliotecas não só transformar e mover seus dados de MARC para BIBFRAME, mas também interligá-los a outros por meio de múltiplos vocabulários, o que contribuirá para a realidade de uma *web* de dados.

Mas, enquanto os estudos para o aprimoramento do BIBFRAME continuam, as bibliotecas podem e devem usufruir das tecnologias *web* disponíveis e ir trabalhando no sentido de agregar significado semântico a alguns elementos de seus registros MARC incluindo as possibilidades abertas pelo LOD.

Nesse contexto, em especial para as bibliotecas brasileiras, uma dessas possibilidades é o uso do formato Dublin Core para a criação de coleções especiais nos repositórios das instituições ou para o desenvolvimento de bibliotecas digitais. Isso é possível porque nos últimos anos, sistemas de gerenciamento de dados, como Pergamum e Sophia, ampliaram suas funcionalidades possibilitando a conversão de um registro MARC em registro DC⁶.

O padrão DC é uma alternativa para o trabalho de enriquecimento semântico dos registros MARC. Desenvolvido pelo grupo de trabalho *Dublin Core Metadata Initiative* (DCMI)⁷ juntamente com o *World Wide Web Consortium* (W3C) o padrão DC utiliza a tecnologia RDF para o processamento de metadados e tem como principal objetivo proporcionar interoperabilidade entre aplicações que trocam informações eletrônicas na *web* (Gracio, 2002).

O padrão de metadados DC se caracteriza não apenas pela simplicidade dos elementos, o que facilita que o criador de um recurso possa descrevê-lo sem a necessidade de extensos treinamentos. Outras características como

⁶ Tais iniciativas alinham-se com a proposta da Biblioteca Digital Brasileira, onde a questão da interoperabilidade é fundamental para o sucesso do projeto, pois possibilita que o usuário consulte de uma única vez todas as fontes de forma integrada, realizando “buscas a recursos informacionais heterogêneos, armazenados em diferentes servidores na rede, utilizando-se de uma interface única sem tomar conhecimento de onde nem como estes recursos estão armazenados” (Marcondes; Sayão, 2001).

⁷ O DCMI é uma organização dedicada à promoção e difusão da adoção de padrões de metadados DC e desenvolvimento de vocabulário especializado para descrever recursos em ambiente web otimizando sua recuperação e acesso (DCMI, 2024).

interoperabilidade semântica, extensibilidade e flexibilidade proporcionam ao usuário liberdade para acrescentar novos elementos e/ou modificá-los para atender uma necessidade de descrição de um determinado recurso, assim como integrá-lo e agregá-lo a outros acervos que também aplicam as tecnologias *web*. Essa integração ocorre porque o DC se tornou um consenso internacional, com o envolvimento no DCMI de mais de vinte países (Gracio, 2002).

As possibilidades de alterar e acrescentar registros e de convertê-los levou Terry Reese a desenvolver o software MarcEdit, ferramenta que pode ser utilizada para editar um registro MARC, incluindo ou alterando campos e subcampos com vistas a agregar significado semântico a alguns de seus elementos, além da possibilidade de convertê-lo para padrões que usam como linguagem de processamento a RDF e incluí-los em uma coleção digital.

3.2 O MARC21 e a Web: editando elementos e agregando objetos digitais

Como vimos, o formato MARC é um esquema de metadados complexo que tem auxiliado as bibliotecas há mais de cinquenta anos no processo de descrição de seus recursos. Além de favorecer a descrição, permite o armazenamento e o compartilhamento dos dados, que podem ser alterados com inclusão, edição e exclusão de campos e subcampos (Mitchell, 2013; 2016).

A disponibilização dos catálogos em linha – *Online Public Access Catalog* (OPAC) – favoreceu a importação de registros, passando este trabalho a ser feito diretamente nos

sítios das bibliotecas. Mas apesar dos protocolos Z39.50 e o *Open Archives Initiative - Protocol for Metadata Harvesting* (OAI-PMH) permitirem o intercâmbio de registros, “[...] os catálogos das bibliotecas [ainda] são ilhas de informação, sem interligação de dados” (Serra; Santarém Segundo, 2017, p. 173). Mas a implementação do *Linked Data* como aplicação prática da Web Semântica pode mudar essa situação.

A proposta do *Linked Open Data* (LOD) está assentada em dois conceitos centrais, onde *linked* define a capacidade de um dado publicado na *web* se conectar com outros dados com informações relacionadas, podendo os mesmos serem compreendidos e acessados por agentes humanos e por computadores. O segundo conceito estabelece que os dados enlaçados para serem utilizados e reutilizados devem ser abertos (*open*), ou seja, livres de restrições de direitos autorais. O mesmo deve ocorrer com vocabulários, ontologias, esquemas etc. que estruturam os conteúdos, cabendo às instituições produtoras desses instrumentos disponibilizá-los colaborando para a construção da Web Semântica (Mitchell, 2013; Serra; Santarém Segundo, 2017).

O LOD oferece grande potencial agregador ao conectar recursos informacionais através de *links* semânticos, *links* que são significativos também para os agentes de *software*. Marcondes (2012) explica: ao contrário dos *links* convencionais, que nada mais são do que etiquetas textuais que levam os programas navegadores a partir de um recurso acessar outro, os “[...] *links* semânticos podem ser processados de forma mais rica por eles, explorando e enriquecendo cognitivamente o significado (legível por máquina) da ligação

entre ambos os recursos” (Marcondes, 2012, p. 173). Isso é possível porque muitas instituições vêm disponibilizando seus acervos através da *web* e estendendo seus serviços para outros usuários que potencialmente podem deles usufruir.

As tecnologias de dados interligados não são tecnologias de recuperação de informações, na acepção que definimos em Ciência da Informação, mas permitem que uma vez identificado um recurso, recuperá-lo e interligá-lo a outros “[...] que lhe agreguem valor semântico, cultural, cognitivo, econômico ou científico” (Marcondes, 2012, p. 174).

Os cuidados com a preservação têm levado os bibliotecários a terem que restringir a circulação de obras, seja porque há poucos exemplares ou são consideradas raras, inibindo assim o acesso. Outro problema enfrentado pelas bibliotecas é a quantidade de recursos escaneados e digitalizados disponíveis na internet.

No sentido de buscar sistematizar o acesso a recursos em acesso aberto ou de domínio público governos e instituições têm sido estimulados a criarem bibliotecas digitais, reunindo conteúdos na forma textual, sonora ou em imagens e que podem ser acessados por qualquer pessoa, como os que estão disponíveis na Biblioteca Domínio Público e na BNDigital.

89

⁸ BNDigital: Biblioteca digital lançada pela Biblioteca Nacional do Brasil em 2006. Integra coleções que desde 2001 vinham sendo digitalizadas no contexto de exposições e de projetos temáticos, em parceria com instituições nacionais e internacionais. (Biblioteca Nacional, 2024).

⁹ Domínio público: biblioteca digital desenvolvida pelo Ministério da Educação do Brasil em 2004. Sua missão é promover o acesso às obras

Outra ferramenta importante é a DBpedia, a maior plataforma de descoberta de conhecimento disponível gratuitamente na *web*. Constituída por meio do recurso *Linked Open Data*, a base extrai conteúdo estruturado das informações da Wikipédia e permite acesso a uma série de informações.

Como vimos, o MARC vem evoluindo junto com as tecnologias de informação sendo o modelo BIBFRAME sua aceção mais avançada, mas ainda em desenvolvimento. Em termos de estrutura, o formato MARC é bem detalhado oferecendo ao catalogador flexibilidade na seleção dos campos a serem utilizados na descrição. Seus campos e subcampos comportam todas as áreas da ISBD, além de outros usados para informações locais, identificados como 9XX e X9X e foram criados para permitir o registro de informações de interesse específico para os usuários locais, “[...] visando a tornar mais completa a descrição bibliográfica e o acesso ao texto completo de um documento” (Oliveira *et al.*, 2004, p. 183). Essas informações podem ser *links* para objetos digitais em acesso aberto (livros, por exemplo) e que estão disponibilizados em Bibliotecas Digitais, como a Europeia. Essa informação pode ser descrita por meio de subcampos que contenham dados sobre a proveniência, a instituição agregadora, o país agregador, o URI do recurso e a base/fonte de localização do recurso. Procedimento menos detalhado, mas que também redundante em enriquecimento semântico é

literárias, artísticas e científicas (na forma de textos, sons, imagens e vídeos), já em domínio público ou que tenham a sua divulgação devidamente autorizada. (Portal Domínio Público, 2024).

adicionar ao registro o campo 590 – notas locais e incluir o URI do recurso.

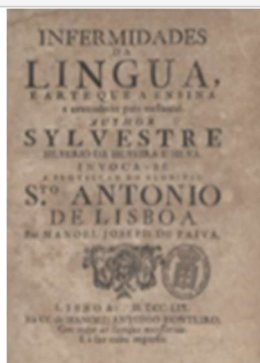
Nesse contexto, o *software* MarcEdit¹⁰ assume a função de instrumento de racionalização de procedimentos, permitindo-nos intervir na estrutura do registro para editá-lo com vista a complementar suas informações e ampliar as possibilidades de acesso a recursos disponíveis na *web*. A ferramenta também permite a tradução de metadados MARC para outros formatos, como o DC, o BIBFRAME, entre outros.

Esse procedimento torna-se mais significativo se o gestor da biblioteca selecionar parte do acervo, usar MarcEdit para acrescentar campos e agregar *links* semânticos e, por fim, converter o registro MARC21 em Dublin Core (DC), cuja linguagem de descrição é a RDF. Assim a biblioteca vai gerando, paralelamente a sua coleção física, uma coleção digital com acesso aos textos. Como exemplo, selecionamos um registro do acervo de obras raras da Biblioteca Nacional do Brasil (Figura 1) e fizemos a edição dos campos MARC21 e, por fim, a tradução para o padrão DC.

¹⁰ Criado em 1999, o MarcEdit começou como um simples utilitário para edição de elementos MARC, mas hoje se tornou um dos mais completos conjuntos de edição de metadados disponíveis para bibliotecários. O programa evoluiu e atualmente inclui uma série de recursos voltados para tornar a tradução e edição de metadados mais fácil para usuários iniciantes e avançados. Esses recursos incluem: edição e integração de dados; adicionar e/ou excluir campos, editar subcampos, trocar dados de campo, copiar dados, realizar a conversão de metadados etc. (Marcedit Development, 2023, tradução nossa).

Figura 1: Registro Marc selecionado

092 __ |a V-087,03,13
100 __ |a Paiva, Manuel José de |d 1706-
245 __ |a *Infermidades da lingua e arte que a ensina a emmudecer para melhorar...*
260 __ |a Lisboa [Portugal] |b na Of. De Manoel Antonio Monteiro |c 1760
300 __ |a 212, [1] p. |c 19 cm.
500 __ |a Data retirada das licencas
500 __ |a Data da pagina de rosto, 1759
500 __ |a Referencia: Innocencio, t.6, p.31
500 __ |a Carimbo: Da Real Bibliotheca
650 __ |a Linguagem e línguas
852 __ |a Livros Raros
990 __ |a Livro

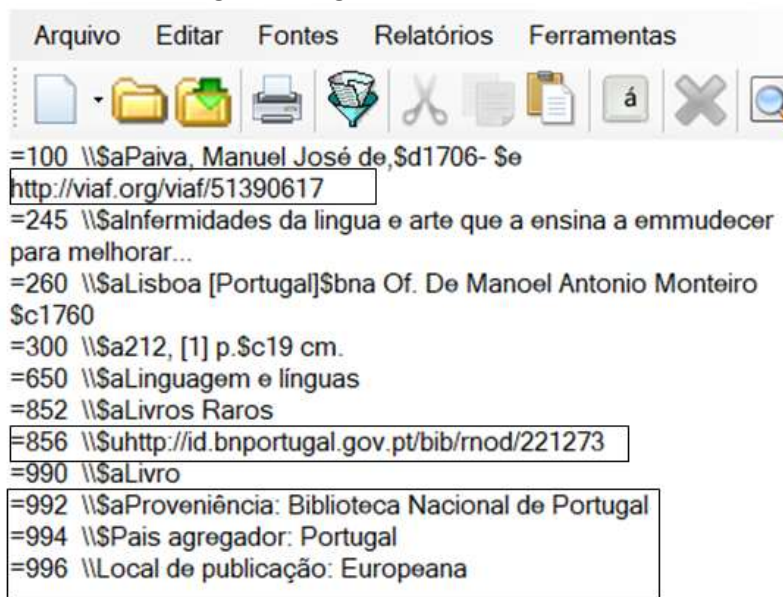


Fonte: Biblioteca Nacional (Brasil).

O registro MARC21 (Figura 1) passou pelo processo de edição, utilizando-se para isso o aplicativo MarcEdit (Figura 2). Alguns campos e subcampos foram subtraídos e outros adicionados de modo a agregar *links* significativos associados às informações como o autor/pessoa (campo 100) e o identificador da publicação na *web*. Um dos campos MARC

adicionados foi o campo 856 já previsto na estrutura do formato para registrar *link* para objetos digitais¹¹.

Figura 2: Registro MARC editado



Fonte: Elaboração das autoras.

Observa-se que no processo de enriquecimento do registro MARC foi agregado um *link* semântico associado ao nome do autor, complementando a informação. Esse *link* permanente (um URI) está disponível no VIAF¹². Em pesquisa

¹¹ Outro campo que pode ser usado pelos gestores das bases de dados é o campo 59X (notas locais), onde agrega-se informações sobre coleção, disponibilização do material etc. Como é um campo repetível, pode ser usado sem restrições. Os campos 590-599 são reservados para número de chamada local e outras definições locais. As informações sobre essas entradas devem ser fornecidas, pelas instituições criadoras da entrada, aos integrantes das redes de intercâmbio (Maranhão; Mendonça, 2017).

¹² O Virtual International Authority File (VIAF) combina vários arquivos de autoridade de nomes num único serviço de autoridade de nomes hospedados pela Online Computer Library Center (OCLC). O objetivo do

na Biblioteca Digital Europeia¹³ encontramos um identificador persistente (um URI) que enlaçava o livro impresso com seu objeto digital. Foi possível fazer essa relação com segurança, sem violar questões de direito autoral, porque o livro encontra-se em acesso aberto. Para indicar as fontes, a título de exemplo, lançamos mão dos campos 9XX para registrar a proveniência do objeto digital, o país agregador e o local de publicação. Também poderíamos usar a DBpedia no campo 500 (notas gerais) ou 59X (notas locais) para agregar informações biográficas sobre o autor ou sobre o editor por se tratar de obra de valor histórico. Ampliando as possibilidades abertas pelo uso do aplicativo MarcEdit poderíamos converter o registro MARC editado (Figura 2) para o formato DC (Figura 3) e depositá-lo em uma coleção do repositório institucional.

serviço é reduzir o custo e aumentar a utilidade dos arquivos de autoridade de bibliotecas, combinando e interligando arquivos de autoridade amplamente usados e disponibilizando essas informações online (VIAF, 2024).

¹³ Europeia é uma biblioteca digital desenvolvida pelos países da União Europeia. Reúne acervos digitais de arquivos, bibliotecas e museus a fim de partilhar o patrimônio cultural para fins de diversão, educação e investigação (Europeana, 2024). Segundo Marcondes (2016), a Europeia é um grande agregador de metadados, coletados (através do protocolo OAIPMH), processados e enriquecidos com novos metadados, gerando um formato interno capaz de representar objetos digitais. A base deste processamento interno e reformatação dos metadados é o modelo conceitual Europeia Data Model (EDM).

Figura 3: Registro MARC traduzido para o padrão DC

```
<rdf:Description>
<dc:title>infermidades da lingua é arte que a ensina a
emmudecer para melhorar /</dc:title>
<dc:creator>Paiva, Manuel José de,1706-</dc:creator>
<dc:creator> http://viaf.org/viaf/51390617</dc:creator>
<dc:publisher>Lisboa : na Of. De Manoel Antonio
Monteiro,</dc:publisher>
<dc:date>1760.</dc:date>
<<dc:description>212 p. ; 19 cm</dc:description>
<dc:description> http://id.bnportugal.gov.pt/bib/rnod/221273
</dc:description>
</rdf:Description> Proveniência: Biblioteca Nacional de
Portugal</dc:description>
</rdf:Description> País agregador: Portugal</dc:description>
</rdf:Description> Local de publicação:
Europeana</dc:description>
```

Fonte: Elaboração das autoras.

O mesmo procedimento poderia ser feito com outras obras raras e paralelamente se criar uma coleção digital. Como os registros DC são descritos em RDF essas informações serão encontradas na internet pelos agentes de *software*. Usufruido da extensibilidade do padrão DC, pode-se lançar mão do elemento *dc:source*, definido como “Uma referência a um recurso de onde o presente recurso possa ter derivado” (Borbinha, 2000, p. 1) para relacionar outras manifestações da obra.

Para a realidade das bibliotecas brasileiras, que sofrem com os poucos recursos financeiros para digitalizar seus acervos, trata-se de uma boa solução para ampliar as possibilidades de acesso aos documentos, muitos deles

exemplares únicos que não podem circular, que são caros ou raros, mas estão disponíveis na *web*, e por meio das tecnologias LOD disponíveis podem ser descobertos e acessados remotamente, complementando os acervos e enriquecendo-os.

No exemplo citado podemos criar outros campos e agregar outras relações culturalmente significativas entre o objeto do acervo e outros existentes em arquivos e museus, disponíveis em acesso aberto na *web*. Nesse caso, apesar de os objetos digitais representarem objetos de acervos diferentes, “que [ainda] não têm equivalência semântica, demonstram ter relações válidas sob um ponto de vista cultural” (Marcondes, 2016, p. 75). Além disso, essas relações se intensificam na medida que objetos de um dado acervo passam a ser representados como objetos digitais, o que nos permite integrá-los com lugares, pessoas, movimentos artísticos, culturais etc.

4 CONCLUSÃO

Em palestra recente Marcondes (2021) chamou a atenção ao afirmar que no Brasil as instituições de cultura e memória pouco aproveitam as tecnologias *web* para enriquecer semanticamente seus acervos. Essa análise levou-nos ao desafio de escrever esse artigo. Trata-se de um artigo com pretensões didáticas de aproximar os profissionais e discentes das questões que envolvem a aplicação das tecnologias *web* nos catálogos, em sua maioria estruturados em MARC21.

A criação do MARC na década de 1960 abriu a possibilidade de se estabelecer um catálogo centralizado que, em fita magnética permitiria de forma mais fácil a consulta a uma grande quantidade de recursos. Atualmente não é diferente com a Web Semântica e suas tecnologias LOD, que buscam interligar dados abertos provocando relações semânticas e culturalmente significativas entre os acervos de instituições de cultura e memória. Procurou-se demonstrar, por meio de exemplos, as possibilidades abertas pelo uso das tecnologias da web nos catálogos estruturados em MARC; a viabilidade de agregar valor aos registros catalográficos complementando-os com dados abertos e interligando-os a outros.

Apesar dos protocolos *web*¹⁴ permitirem o intercâmbio de registros MARC, os catálogos das bibliotecas são silos de dados sem interligação. A implementação de *Linked Data* como aplicação prática da Web Semântica pode mudar essa situação. O LOD e outras tecnologias *web* e as ações correlatas a elas, como linguagem/vocabulário RDF e os modelos conceituais, como o *Europeana Data Model (EDM)*, ampliam as possibilidades de representar objetos digitais, retirando-os dos estreitos limites dos catálogos tradicionais tornando-os, interoperáveis e acessíveis, tanto para os humanos como para as máquinas.

O esforço em tornar os catálogos interoperáveis e acessíveis às pesquisas por humanos e máquinas vem exigindo da *Library of Congress* o aprimoramento do formato MARC

¹⁴ Protocolos Z39.50 e *Open Archives Initiative-Protocol for Metadata Harvesting (OAI-PMH)*.

com vista a termos uma estrutura de metadados que abarque a complexidade das descrições documentárias e que possa expressar-se de forma interoperável. Busca-se com as marcações proporcionadas pela linguagem RDF elaborar descrições por meio de triplas passíveis de revelar a natureza das entidades e seus atributos, além de permitir relacionamentos e enlaçamentos de dados. Foi com o propósito de aproximar o velho formato – o MARC - das tecnologias *web* que os projetos MarcOnt e BIBFRAME foram lançados, esse último o mais completo e robusto, capaz de suportar representações descritivas mais detalhadas, que abrigam muitas notas locais, já que o compartilhamento de registros entre unidades de informação de um mesmo sistema ou comunidade é uma estratégia de gestão para otimizar o processo de catalogação e economizar recursos.

O uso simultâneo de diversos vocabulários por meio de ligação de dados tem sido uma das mais importantes estratégias do *Linked Data*, a qual contribui para enriquecer semanticamente as tarefas realizadas na *web*. Quando aos registros MARC essas ações são incorporadas é possível ampliar as possibilidades de as bibliotecas atenderem as tarefas de seus usuários, conforme preconiza os Princípios da Catalogação. Essa percepção está presente no modelo BIBFRAME que já estuda a possibilidade de reutilizar vocabulários devido a reconhecer seus benefícios para o trabalho de representação. No entanto, essa é uma questão delicada e que requer cuidadosa consideração, já que diversos fatores podem afetar o desenvolvimento de sistemas, como

fatores econômicos, mudanças organizacionais, interesses sociais etc.

No Brasil o fator econômico é o que mais afeta o desenvolvimento de nossas bibliotecas. Estratégias criativas como a proporcionada pela edição do registro MARC, com vista à inclusão de LOD de objeto digital correspondente ou dele derivado (novas expressões ou manifestações da obra), assim como, a tradução de seus metadados em padrão compatível com as tecnologias web, como o padrão DC, são possibilidades que se abrem para o enriquecimento dos nossos catálogos, ampliando para os usuários e sociedade em geral o acesso aos acervos e aos seus conteúdos.

REFERÊNCIAS

BERNERS-LEE, T. *Linked data: design issues*. 2006. Disponível em: <https://www.w3.org/DesignIssues/LinkedData.html>. Acesso em: 31 maio 2021.

BERNERS-LEE, T.; HENDLER, J.; LASSILA, O. The semantic web. *The Semantic Web*, [s. l.], v. 284, n. 5, p. 28–37, maio 2001.

BIBLIOTECA NACIONAL (Brasil). *Sobre a BN Digital*. Rio de Janeiro: FBN, 2024. Disponível em: <https://bndigital.bn.gov.br/sobre-a-bndigital/>. Acesso em: 06 dez. 2024.

BIZER, C.; HEATH, T.; BERNERS-LEE, T. Linked data: the story so far. In: HEATH, T.; HEPP, M.; BIZER, C. (Ed.). *International Journal on Semantic Web and Information System*, [S. l.], v. 5, n. 3, p. 1-22, 2009. Special Issue on Linked Data.

BORBINHA, José Luis. *DCMI: elementos do núcleo de metadados Dublin Core, versão 1.1*, [Lisboa]: Biblioteca Nacional, 2000.

CASTRO, F. F. de; SANTOS, P. L. V. A. da C. MarcOnt Initiative: representação e descrição de recursos informacionais na Web. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 9., 2008, São Paulo. *Anais [...] São Paulo: USP, 2008*. Disponível em: <http://repositorios.questoesemrede.uff.br/repositorios/bitstream/handle/123456789/1896/Marcont.pdf?sequence=1> Acesso: 8 jul. 2021.

DCMI. *Dublin Core Metadata Initiative*. 2024. Disponível em: <https://www.dublincore.org/>. Acesso: 6 dez. 2024.

EUROPEANA Discover Europe's digital cultural heritage. 2024. Disponível em: <https://www.europeana.eu/en>. Acesso em: 13 dez. 2024.

FIUZA, M. M. A catalogação bibliográfica até o advento das novas tecnologias. *Revista da Escola de Biblioteconomia da UFMG*, Belo Horizonte, v. 16, n. 1, p. 43-53, 1987.

GRÁCIO, J. C. A. *Metadados para a descrição de recursos da Internet: o padrão Dublin Core, aplicações e a questão da interoperabilidade*. 2002. 127 f. Dissertação (mestrado) - Universidade Estadual Paulista, Faculdade de Filosofia e Ciências, 2002. Disponível em: <http://hdl.handle.net/11449/93722>. Acesso em: 10 jul. 2021.

HEATH, T.; BIZER, C. *Linked data: evolving the web into a global data space*. San Rafael, USA: Morgan & Claypool Publishers, 2011.

IFLA. *Declaração dos Princípios Internacionais de Catalogação (PIC)*. Haia: IFLA, 2016. Disponível em: [50](https://www.ifla.org/wp-</p></div><div data-bbox=)

content/uploads/2019/05/assets/cataloguing/icp/icp_2016-pt.pdf. Acesso em: 18 jul. 2021.

LIBRARY OF CONGRESS. BIBFRAME. *Bibliographic framework initiative*. Disponível em: <https://www.loc.gov/bibframe/>. Acesso em: 4 jul. 2021.

MARANHÃO, A. M., MENDONÇA, M. de L. dos S. *MARC 21 - formato bibliográfico*. Rio de Janeiro: PUC-RJ, 2017.

MARCEdit DEVELOPMENT. *About MarcEdit*. 2013. Disponível em: <https://marcedit.reeset.net/about-marcedit> Acesso: 15 Jul. 2021

MARCONDES, C. H. "Linked Data"- dados interligados e interoperabilidade entre arquivos, bibliotecas e museus na web. *Encontros Bibli: revista eletrônica de biblioteconomia e ciência da informação*, Florianópolis, v. 17, n. 34, p. 171-192, maio/ago., 2012.

MARCONDES, C. H. Interoperabilidade entre acervos digitais de arquivos, bibliotecas e museus: potencialidades das tecnologias de dados abertos interligados. *Perspectivas em Ciência da Informação*, Belo Horizonte, v.21, n.2, p.61-83, abr./jun. 2016.

MARCONDES, C. H. Representação e economia da informação. *Ciência da informação*, Brasília, v. 30, n. 1, p. 61-70, jan./abr. 2001.

MARCONDES, C. H. *Tecnologias de dados abertos interligados e as bibliotecas*. Belo Horizonte: UFMG/PPG-GOC, 2021.

MARCONDES, C. H.; SAYÃO, L. F. Integração e interoperabilidade no acesso a recursos informacionais eletrônicos em C&T: a proposta da Biblioteca Digital Brasileira. *Ciência da Informação on-line*, Brasília, v. 30, n. 3, p. 24-33, 2001.

MEY, E. S. A. *Não brigue com a catalogação!* Brasília: Briquet de Lemos/Livros, 2003.

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

MITCHELL, E. T. *Library linked data early ctivity and development*, Chicago: ALA TechSource, 2016.

MITCHELL, E. T. *Library linked data: research and adoption*. Chicago: ALA TechSource, 2013.

OLIVEIRA, Z. P. de *et al.* O uso do campo MARC 9XX para controle bibliográfico institucional. *Ciência da informação*, Brasília, v. 33, n. 2, p. 179-186, maio/ago. 2004. Disponível em: <http://revista.ibict.br/ciinf/article/view/1060/1145> Acesso em: 10 jul. 2021.

PORTAL DOMÍNIO PÚBLICO. *Missão da Biblioteca Virtual Domínio Público*. Brasília: MEC, 2024. Disponível em: <http://www.dominiopublico.gov.br/Missao/Missao.jsp>. Acesso em: 06 dez. 2024.

SERRA, L. G.; SANTARÉM SEGUNDO, J. E. O catálogo da biblioteca e o linked data. *Em Questão*, Porto Alegre, v. 23, n. 2, p. 167-185, maio/ago. 2017.

SILVEIRA, N. C.; TÁLAMO, M. de F. G. M. Os FRBR e a escolha do ponto de acesso pessoal. *Perspectivas em Ciência da Informação*, João Pessoa, v. 14, n. 2, p. 108-120, maio/ago. 2009.

SOUZA, R. R.; ALVARENGA; L. A Web semântica e suas contribuições para a ciência da informação. *Ciência da Informação*, Brasília, v. 33, n. 1, p. 132-141, jan./abr. 2004.

VIAF. *The virtual international authority file*. 2024. Disponível em: <https://viaf.org/>. Acesso: 06 dez. 2024.

YANKULOV, M. *What is SPARQL?* New York: ONTOTEXT, [2024]. Disponível em: https://www.ontotext.com/knowledgehub/fundamentals/what-is-sparql/?generate_pdf=13611. Acesso em: 13 dez. 2024.

ZAFALON, Z. R.; SANTOS, P. L. V. A. da C. Sintaxe e semântica de registros bibliográficos: princípios para a conversão de registros analógicos para o formato MARC21 bibliográfico: Scan for MARC. *Tendências da Pesquisa Brasileira em Ciência da Informação*, [São Paulo], v. 5, n. 1, p. 1-23, 2012. Disponível em: <http://hdl.handle.net/11449/114826>. Acesso em: 28 maio 2021.

Z Aidan, F. H.; Bax, M. P. Linked open data como forma de agregar valor às informações clínicas. *AtoZ: novas práticas em informação e conhecimento*, Curitiba, v. 2, n. 1, p. 44-59. jan./jun. 2013. Disponível em: <http://www.atoz.ufpr.br>. Acesso em: 28 maio 2021.

CAPÍTULO 2

METADADOS PARA REPRESENTAÇÃO DE CONJUNTOS DE DADOS DE PESQUISA E OS PRINCÍPIOS FAIR

Caterina Marta Groposo Pavão

Cleusa Pavan

1 INTRODUÇÃO

O compartilhamento de dados de pesquisa tem crescido como efeito da demanda das agências de financiamento públicas e privadas e dos periódicos científicos, para os cientistas com projetos de pesquisa aprovados e no processo de *peer review*, respectivamente. Porém, mais do que cumprir com requerimentos, ao disponibilizar dados de pesquisa em acesso aberto, na internet, o cientista beneficia a sua área do conhecimento como um todo, encorajando os pares a realizarem novas investigações e descobertas. Sob o aspecto individual e de sua carreira, ele usufrui de uma plataforma própria, os repositórios, para organizar, preservar e disponibilizar dados de forma imediata, amplia a visibilidade de suas investigações e recebe crédito por meio da citação.

Esse novo processo está no escopo da Ciência Aberta que tem como princípio básico a convicção de que o conhecimento é um bem público e que grande parte das pesquisas científicas é desenvolvida com fomento governamental, devendo ficar mais acessível para a

comunidade científica, o setor de negócios e a sociedade em geral. Segundo o documento de recomendações da Organização das Nações Unidas para a Educação, a Ciência e a Cultura (UNESCO) sobre Ciência Aberta, publicado em 2020, as soluções sustentáveis e inovadoras requerem esforços científicos eficientes, transparentes e dinâmicos, não somente da comunidade científica, mas também de toda a sociedade. Para assegurar que a ciência realmente favoreça as pessoas e o planeta, é necessário transformar todo o processo científico. Sendo assim, a Ciência Aberta é o movimento que pretende essa transformação, disponibilizando os dados, a informação e a produção científica ao alcance de todos, de forma democrática e confiável.

Esse movimento engloba uma série de práticas, além dos dados de investigação, como softwares e ferramentas de código aberto, fluxos de trabalho abertos, o acesso aberto às publicações, ciência cidadã, recursos educativos abertos, ciência aberta reproduzível, políticas para ciência aberta e métodos alternativos para avaliar a investigação, incluindo a revisão pelos pares aberta (Pontika *et al.*, 2015). Nesta estrutura da Ciência Aberta, os dados de pesquisa ocupam um lugar extremamente relevante no momento em que cada vez mais os recursos são escassos e partilhar conhecimento é a palavra de ordem. O compartilhamento de dados de pesquisa é basicamente o ato de disponibilizá-los para sua reutilização.

Os dados da pesquisa abrangem todas as disciplinas e, portanto, sua definição ou conceito pode variar em relação às abordagens dos diferentes atores (pesquisadores, instituições, órgãos de fomento, governos, entre outros) e dos diferentes

contextos nacionais. Resumidamente, podemos dizer que são as informações registradas ou produzidas por qualquer forma ou meio durante o processo de investigação. Em 2007, a *Organization for Economic Co-operation and Development* (OECD) publicou um conjunto de diretrizes com base em princípios comumente aceitos para facilitar o acesso otimizado e econômico aos dados de pesquisa digital financiados com recursos públicos,

No contexto destes princípios e diretrizes, os dados de pesquisa são definidos como registros factuais (registros numéricos, textuais, imagens e sons) usados como fontes primárias para pesquisa científica, e que são comumente aceitos na comunidade científica como necessários para validar resultados da pesquisa. Um conjunto de dados de pesquisa constitui uma representação sistemática e parcial do assunto que está sendo investigado. (OECD, 2007, p. 13, tradução nossa)

Chapman *et al.* (2020) definem um conjunto de dados como um conjunto de imagens, gráficos ou documentos, bem como a tabela clássica de dados, relacionadas, organizadas e formatadas para um propósito específico. Os autores, ainda, colocam que a busca de conjuntos de dados envolve a descoberta, a exploração e o retorno de conjuntos de dados, como resultado, a um usuário final. Segundo Löffler *et al.* (2021), esta definição é um pouco restritiva, pois todos os tipos de dados científicos, como os experimentais, observacionais, ambientais e genômicos, simulações e dados computacionais podem ser considerados como conjuntos de dados. Assim, fazem uma releitura da definição de Chapman e colegas, da seguinte forma: “[...] um conjunto de dados é uma coleção de

dados científicos, incluindo dados primários e metadados organizados e formatados para um propósito específico” (Löffler *et al.*, 2021, p. 5, tradução nossa).

Ainda, conjuntos de dados de pesquisa podem ser considerados recursos com atributos que devem ser descritos por meio de estruturas de representação, que permitam sua busca e recuperação em um determinado sistema. A sistematização de dados de pesquisa é um processo que consiste em organizar os dados por meio de normas e padrões interoperáveis, que permitem sua recuperação e consequente acesso pela comunidade científica e acadêmica. Assim, os dados da pesquisa devem ser organizados para encorajar uma possível reutilização pelos sistemas abertos (Ávila Barrientos, 2020).

Os dados de pesquisa são o resultado dos projetos de investigação, usados como fontes primárias que sustentam a investigação científica e possibilitam a derivação de descobertas teóricas ou aplicadas. Devem, fundamentalmente, tornar as descobertas ou os estudos replicáveis, ou pelo menos reprodutíveis, ou precisam ser, em alguma medida, reutilizáveis. Para tal, Sales e Sayão (2019), colocam que os dados devem ser documentados, ter proveniência comprovada, estar armazenados em lugar seguro onde a preservação seja garantida para usos atuais e futuros, ser compartilhados, acessados e reusados. A fim de atender estas premissas, as boas práticas orientam para que os dados sejam o mais abertos e FAIR possíveis, mas sem esquecer as questões éticas, comerciais e de privacidade de dados sensíveis ou dados proprietários.

Os princípios FAIR surgiram em 2014, com o objetivo de apresentar normas e boas práticas para facilitar a reutilização dos dados de pesquisa. A sigla identifica as quatro características que os dados ou os objetos digitais devem possuir: *Findable* (localizáveis), *Accessible* (acessíveis), *Interoperable* (interoperáveis) e *Reusable* (reutilizáveis). De acordo com Rocha *et al.* (2019), para os dados serem qualificados como FAIR, a estes devem ser atribuídos identificadores únicos, persistentes e globais (**F**). Devem ser descritos por metadados indexáveis e ricos (**F**), representados em linguagens formais (**I**), aceitos pela comunidade (**R**), com atributos relevantes, precisos e úteis ao contexto (**R**), incluindo metadados de proveniência (**R**) e usando vocabulários controlados que seguem princípios FAIR (**I**). Estes dados devem ser recuperáveis pelo seu identificador, através de um protocolo de comunicação padronizado, aberto, gratuito (**A**). Também, devem ser acompanhados de licenças claras e acessíveis (**R**), e referências qualificadas, devem ligar (meta)dados para enriquecer o conhecimento sobre os mesmos (**I**). Os princípios FAIR não são um padrão, não ditam uma especificação tecnológica ou uma solução para implementação, mas um guia para assistir produtores e publicadores de dados em suas escolhas na gestão dos mesmos, enfatizando o aumento da capacidade das máquinas encontrarem e utilizarem automaticamente os dados, para além do reuso pelos indivíduos (Wilkinson *et al.*, 2016).

O objetivo deste capítulo é discorrer sobre a representação dos dados de pesquisa, com o intuito de torná-los localizáveis, acessíveis, interoperáveis e reutilizáveis.

Utilizar metadados completos e precisos é importante em vários processos, como o reuso e compartilhamento de conjuntos de dados entre os cientistas, a aplicação de curadoria digital e estratégias de proveniência dos dados, e a análise do conteúdo dos repositórios (Rousidis *et al.*, 2014). Os princípios FAIR são mínimos e devem ser observados especialmente pelos produtores de dados, que executam o autoarquivamento de conjuntos ou pacotes de dados (*datasets*, em inglês), e pelos mantenedores de repositórios. Neste capítulo, abordamos a representação da informação no contexto dos repositórios de dados. Para tal, trataremos questões relativas ao emprego de metadados para descrição e a sua qualidade, padrões de metadados internacionalmente aceitos pelas áreas do conhecimento, os valores contidos nos metadados, o cientista produtor, depositante e catalogador de dados, em repositórios de dados de pesquisa, e o papel do profissional da informação.

2 REPOSITÓRIOS DE DADOS DE PESQUISA

Um repositório é um serviço de informação dedicado ao gerenciamento, reunião, armazenamento, organização, preservação, recuperação e ampla disseminação da informação produzida em uma instituição por seus pesquisadores. Constitui-se em uma coleção de objetos digitais que consiste, basicamente, em um banco de dados, no qual se armazena o conteúdo em formato digital, em um servidor web, com um conjunto de interfaces e funções para envio de conteúdo, gerenciamento de sistema, pesquisa e recuperação de conteúdo e outros serviços para o usuário. Os

repositórios digitais devem refletir o conceito de "ambiente de confiança", que foi promovido durante a *International Conference 2012 Cultural Heritage online - Trusted Digital Repositories*, desenvolvido a partir de requisitos, normas, recomendações e diretrizes confiáveis e amplamente aceitas.

Um repositório digital, para ser considerado confiável, deve ter como missão fornecer acesso de longo prazo aos recursos digitais por ele gerenciados e, portanto, aceita a responsabilidade pela manutenção, também a longo prazo, desses recursos, em nome de seus depositantes e em benefício dos usuários atuais e futuros. Para que esse propósito seja atingido, devem seguir o *Open Archival Information System (OAIS)*, um padrão de referência para o desenvolvimento de repositórios que estão comprometidos em preservar e manter seus conteúdos digitais acessíveis a longo prazo. Ele especifica o ambiente do repositório, a forma de representação das informações a serem preservadas e o modelo funcional do repositório. Também, orienta na identificação dos produtores, do processo de produção dos dados, das comunidades-alvo (consumidores) e dos encarregados de sua gestão, especifica as entidades funcionais que compõem um repositório: ingestão, acesso, planejamento da preservação, armazenamento, gestão de dados e administração, serve para delimitar, caracterizar e orientar os componentes de funcionamento do repositório.

Os repositórios de dados exercem um papel fundamental na ciência. Implementam práticas sistemáticas de gestão, incluindo a curadoria, preservação, disponibilidade a longo prazo, disseminação e acesso. São amplamente

difundidos nas comunidades que os produzem, como na Física (SDSS - FermiLab, na Genética (GenBank na NCBI Data) ou nas Ciências ambientais (CEDA Archive).

Para cumprir tais papéis, os repositórios de dados, por meio da tecnologia, devem garantir a descoberta dos dados ao longo do tempo, mas para atender o quarto princípio FAIR, dependem, pelo menos parcialmente, do correto uso dos metadados e da quantidade dos mesmos, para que a descrição de dados reflita as necessidades de informações acadêmicas. Se a informação não está disponível nos metadados, os dados não podem ser recuperados, reutilizados e citados. O estudo de Löffler *et al.* (2021), na área de Biodiversidade, revela que grandes repositórios de dados seguem apenas parcialmente a recomendação do quarto princípio FAIR. Os repositórios de dados gerais tendem a usar apenas os metadados padrão, embora existam diretrizes e padrões de dados de metadados adequados. Assim, os autores se fazem os seguintes questionamentos: por que os padrões adequados às áreas, com as suas especificidades são tão raramente usados? As recomendações são muito amplas? Os padrões contêm muitos campos e elementos? Precisa-se melhorar as ferramentas de software e métodos para facilitar a integração de diferentes metadados? Os autores acreditam que apenas mais estudos de usuários e discussões na comunidade de pesquisa podem ajudar a encontrar respostas para essas perguntas.

Os repositórios de dados, segundo Monteiro e Sant'Ana (2018), têm sua gênese com a necessidade de gestão dos dados científicos, estão vinculados às universidades e instituições de pesquisa e contribuem para assegurar que os

dados sejam publicados e disponibilizados para a comunidade científica com o menor número possível de restrições. Podemos encontrar diversos desafios que tornam difícil descobrir, compartilhar ou reutilizar dados, como os metadados incompletos, imprecisos e/ou mal descritos. Outras barreiras são a baixa visibilidade, problemas com acessibilidade (acesso restrito, contato necessário com o produtor) e usabilidade limitada, devido aos metadados, como ausência de parâmetros-chave para avaliar a proveniência e qualidade dos dados, e para identificar a diversidade da amostra da população (quando for o caso) (Khan *et al.*, 2021). Os autores, que são vinculados às Ciências da Saúde, demonstram receio de os usuários secundários utilizarem *datasets* problemáticos, sem metadados básicos que informem aspectos importantíssimos na pesquisa clínica, sobre a população, como idade, sexo, raça, pacientes grávidas, além de amostras pequenas, pelo simples fato desses dados estarem disponíveis e terem alta visibilidade e fácil acesso no momento da busca em repositórios.

Para Ávila Barrientos (2020), o principal desafio é convencer as instituições a participarem do modelo de acesso aberto, implementando políticas de promoção de dados abertos. O estabelecimento de políticas institucionais norteia e incentiva os pesquisadores para uma melhor administração dos dados, inclusive dos conjuntos coletados há mais tempo e que ainda têm utilidade para a comunidade da área, podendo ser compartilhados tardiamente.

Com a consciência crescente em relação à citação de dados e que, para isso, os conjuntos de dados precisam ser

preservados, descritos e acessados corretamente, as atividades relacionadas à curadoria de dados tornam-se cada vez mais importantes. A fim de que um conjunto de dados seja citado, ele deve primeiro ser arquivado em um repositório, preservado em uma forma interoperável, adequadamente descrito por um conjunto formal de metadados conectados ao conjunto de dados e disponibilizados para reutilização (Corrêa, 2016). Vários padrões de metadados de metadados integral tal conjunto formal. Em alguns casos descrevem o conteúdo de arquivos, em outros as características técnicas dos arquivos e em outros expressam as relações entre os arquivos de um conjunto de dados. Ainda, alguns padrões de metadados servem como suporte para estabelecer uma estratégia de metadados capaz de descrever seus dados e atender às suas necessidades de gerenciamento.

3 METADADOS PARA REPRESENTAÇÃO DE DADOS DE PESQUISA

Na opinião de Solodovnik (2013), não importa a prática de gestão de informação escolhida pela instituição para processar dados bibliográficos, a qualidade dos mesmos depende da sua adequação aos requisitos das normas bibliográficas, bem como, os códigos de catalogação e regras de classificação e vocabulários controlados, utilizados não apenas para definir e descrever registros formalmente, mas também para normalizar seus dados, atributos e valores e, assim, criar um acesso controlado e uniforme de recuperação, interoperabilidade e reutilização de dados bibliográficos. Se bem esta citação refere-se a dados bibliográficos, os princípios

apresentados aplicam-se a outros tipos de conteúdo, documentos ou objetos, sejam eles digitais ou não.

A catalogação tradicional organiza todo um conjunto de dados: autor, título, data, assunto, entre outros, armazenados em um sistema que contém as coleções de uma biblioteca. Os metadados correspondentes a estes dados são considerados componentes lógicos centrais do sistema utilizado em bibliotecas e repositórios digitais. Segundo Campbell (2006), com o surgimento de novos sistemas de metadados, que se encontram nos próprios documentos e afetam a maneira como eles são usados, a catalogação mudou-se para um novo território difícil, em que nosso papel tradicional como intermediários de informação tornou-se menos rigidamente definido e mais sutilmente poderoso.

O esquema de metadados que se pretende utilizar em um repositório digital, seja de documentos ou de dados, institucional ou temático, deve determinar os requisitos de descrição de cada elemento, promover a padronização, normalização e enriquecimento dos metadados para fortalecer a qualidade dos registros e da informação contida nos objetos digitais. Antes de disponibilizar um objeto digital, deve-se ter certeza de que todos os metadados necessários foram registrados. Ainda, no momento do planejamento de um repositório, deve-se avaliar se um único esquema de metadados atende a natureza de todas as coleções ou se será necessário combinar esquemas para melhor atender às necessidades de descrição (Pavão *et al.*, 2015).

Da perspectiva dos dados, os metadados são informações estruturadas e descritivas, criados para

finalidades diferentes, como pesquisa, classificação ou formação de conhecimento e devem dar resposta para: o que foi medido/coletado/pesquisado? por quem? quando? onde? e por quê? Sales e Sayão (2019) argumentam que, para os dados serem encontrados, identificados, compreendidos, reusados e tenham sua proveniência comprovada, eles devem estar acompanhados de um conjunto de informações que descreva todos os seus aspectos, como ilustra a Figura 1.

Figura 1: Conteúdo da descrição dos dados



Fonte: Sales e Sayão (2019).

Esquemas ou padrões de metadados podem ser de amplo espectro, *Dublin Core* e *DataCite*, destacados abaixo, ou voltados a um domínio, como, por exemplo: *Darwin Core* (Biodiversidade), *Data Documentation Initiative* (Ciências sociais e econômicas), *Ecological Metadata Language* (Ecologia e Ciências ambientais e da terra), *EngMeta* (Engenharia computacional) e *ISO 19115* (Ciências geoespaciais).

Dublin Core Metadata Initiative: criado na década de 1990, é um padrão simples e flexível, de fácil uso por não catalogadores, com 15 elementos básicos para descrição de

recursos/objetos na internet, (Colaborador, Cobertura, Criador, Data, Descrição, Formato, Identificador, Língua, Publicador, Relação, Direitos, Recurso, Assunto, Título e Tipo).

DataCite Metadata Schema Documentation: esquema genérico para descoberta e citação de dados, desenvolvido pelo *DataCite Consortium*, mas que possibilita a customização para atender às disciplinas. Apresenta três níveis de obrigação: mandatório, recomendado e opcional. As propriedades obrigatórias são: Identificador, Criador, Título, Publicador, Ano de publicação, Tipo de recurso, enquanto Assunto, Colaborador, Data, Identificador relacionado, Descrição e Geolocalização são recomendadas (DATA CITE, 2021).

Para selecionar os elementos de metadados que serão utilizados, deve-se fazer o confronto contínuo com os seis princípios dos “bons metadados” publicados pela National Information Standards Organization (NISO) (2007), quais sejam:

- 1) devem estar em conformidade com os padrões da comunidade, de forma apropriada para os materiais da coleção, os usuários da coleção e para os atuais e potenciais futuros usuários;
- 2) oferecer suporte à interoperabilidade;
- 3) usar controle de autoridades e de conteúdo para descrever e relacionar objetos;
- 4) incluir uma declaração clara das condições e termos de uso para o objeto digital;
- 5) oferecer preservação e curadoria digital a longo prazo dos objetos nas coleções;

6) são objetos em si e, portanto, devem ter as qualidades de bons objetos, incluindo autoridade, autenticidade, arquivamento, persistência e identificador único.

Podemos verificar que os princípios FAIR apresentam características que se sobrepõem àqueles da NISO, pois consideraram os metadados um componente-chave para o compartilhamento de *datasets*, sendo valorados tanto quanto os dados em si. Somente uma descrição normalizada leva à recuperação e reuso dos conjuntos de dados num repositório de forma mais eficiente.

Para Solodovnik (2013), a validação da qualidade de metadados é uma questão de gestão organizacional em vez de um procedimento. Portanto, é uma boa prática estabelecer uma unidade de apoio dirigida por profissionais de metadados dentro de cada repositório. A redação de instruções e recomendações são essenciais para apoio ao produtor de dados.

Tais profissionais devem buscar o aprimoramento dos metadados de descrição de conteúdo, considerando a missão de manter um padrão de alta qualidade, em conformidade com as mais recentes tecnologias, e desenvolver recursos, a fim de que os conteúdos possam ser inequivocamente identificados, tendo em vista a sua disseminação. Wu *et al.* (2019) elegem dez recomendações para os repositórios aperfeiçoarem a descoberta de dados e a experiência dos usuários no momento da busca. Entre elas, tornar os registros de metadados legíveis e analisáveis, sendo que os campos mais importantes deveriam ficar no topo das páginas, nomeados

devidamente para não gerar ambiguidades, com *hyperlinks* e botões facilmente reconhecíveis; identificar e agregar registros de metadados de um mesmo *dataset*, uma vez que podem ser disponibilizados em outros repositórios, de forma diferente, quando fruto de pesquisa em colaboração entre instituições; e facilitar a indexação e busca dos registros de metadados pelos motores de busca de *datasets*, empregando ferramentas existentes para tal.

É preciso ressaltar que, em oposição aos catálogos de bibliotecas, bases de dados e outros serviços especializados de indexação, em que o profissional da informação é o principal responsável pela descrição dos objetos, nos repositórios de dados de pesquisa a atividade é majoritariamente do cientista. White (2014) argumenta que os dados de pesquisa são um tipo de informação pessoal, organizada de forma pessoal, fatos que precisam ser considerados antes de criar sistemas de informação e integrar qualquer um dos esquemas de metadados já estabelecidos. A coleta e organização de dados, para uso próprio e de seus colaboradores, é uma ação básica do fazer científico diário. Entretanto, organizar com a finalidade de compartilhar os dados, de forma ampla e global, somente tornou-se um requerimento, mais recentemente, para a maioria dos pesquisadores. Isso implica alterar os procedimentos, até então praticados, e dedicar mais tempo à descrição dos *datasets*, com ênfase na descoberta e compartilhamento e não apenas na análise para publicação nos tradicionais canais de comunicação científica, como artigos, livros, relatórios; e à orientação dos integrantes de seu

grupo de pesquisa na coleta e imediata documentação dos dados.

Ao comparar profissionais da informação e pesquisadores, White (2014) observa que ambos fazem uso de diretrizes, sejam padrões formais estabelecidos ou guias de uso individual ou local, sendo que os primeiros possuem mais experiência na utilização de metadados, e os pesquisadores são mais específicos, menos focados na descrição, e mais preocupados com os *softwares* empregados. Para a autora, fica assinalada a distinção no papel do cientista como depositante e do profissional da informação como curador, um aspecto que interfere na arquitetura dos repositórios. As diferenças são esperadas, pela bagagem profissional de cada um dos grupos. Cabe aos mantenedores dos repositórios mapear antecipadamente as necessidades das comunidades-alvo, e conciliar uma equipe de trabalho para estudo e acompanhamento do(s) esquema(s) de metadados adotado(s).

Rousidis *et al.* (2014), em estudo no repositório Dryad, validaram a suposição dos problemas de qualidade ocorrer, majoritariamente, com os elementos Criador, Data e Tipo. Quando focam no campo Assunto, detectam média de 4,79 palavras-chave por pacote de dados, e várias questões de qualidade, como erros de digitação, entrada de informação irrelevante, extrema diversidade de sinônimos, inconsistência entre plural e singular (Rousidis *et al.*, 2015). Os autores sugerem um maior controle manual da entrada de dados, uma das etapas da curadoria.

Em outro estudo, com *datasets* de imagens oftalmológicas, Khan *et al.* (2021) analisam 94 conjuntos de dados em acesso aberto (507.724 imagens, 125 vídeos) e verificam que detalhes técnicos das imagens e sua aquisição foram registrados, porém não ocorreu o mesmo com informação clínica. As características dos pacientes (sexo, idade, origem étnica) constavam em menos de 20% dos *datasets*, o período de coleta em 19% e os critérios de inclusão e exclusão em apenas 15%. Nota-se descuido do depositante com elementos descritivos indubitavelmente reconhecidos por ele, no papel de pesquisador, como essenciais no desenho de qualquer pesquisa, coleta de dados e relato dos resultados. O esperado é que os cientistas mantenham o mesmo zelo ao compartilhar os dados no repositório e ao realizar a pesquisa.

Certamente os padrões melhoram o fluxo de informações dentro das comunidades científicas, mas também podem criar barreiras e impedir a inovação quando são prematuros ou inadequados (Borgman, 2015). Mais uma vez, os mantenedores de repositórios e profissionais da informação têm papel fundamental em avaliar o nível de desenvolvimento do esquema a ser empregado, a necessidade de inclusão de elementos novos que deem conta de uma especificidade institucional e a junção de esquemas especializados a um genérico. A execução de testes-piloto com os cientistas poderá evitar incômodos futuros e retrabalho.

4 OS METADADOS E O PROFISSIONAL DA INFORMAÇÃO

O profissional da informação tem expandido o escopo das suas atividades para se adequar às mudanças tecnológicas e às demandas dos usuários por novos serviços, antes mesmo do surgimento das bibliotecas digitais, logo em seguida com os repositórios institucionais e, mais recentemente, os repositórios de dados de pesquisa, gerando novos significados, revisando conceitos antigos e vendo seu espaço profissional crescer, absorvendo novas funções e desenvolvendo competências.

Esse profissional está capacitado para atuar em ambientes onde o ciclo da informação acontece, desde a seleção e aquisição, passando pela representação descritiva e temática, pelo armazenamento, até a descoberta e disseminação da informação. “Não é o diploma que concede o papel a esse profissional, mas é o mercado o regulador.” (Boeres, 2017, p. 36). Para a mesma autora, o profissional da informação é fruto da sua formação, que congrega conhecimentos da Ciência da Informação, uma Ciência social aplicada, que tem como objetivo estudar todos os aspectos da informação e analisar os processos da sua construção, da sua comunicação e do seu uso. Portanto, o profissional da informação necessita buscar capacitação e atualização constantemente para acompanhar os novos tempos e suas demandas. Foi no final da década de 1990 que surge o chamado Bibliotecário de metadados, refletindo a mudança do papel do Bibliotecário de catalogação.

Em estudo realizado em 2010, Han e Hswe buscaram determinar as competências dos bibliotecários de metadados em comparação com as dos bibliotecários de catalogação. Os dados coletados mostram que o conjunto de habilidades necessárias para um bibliotecário de metadados estão relacionadas com: o conhecimento de padrões de metadados (incluindo MARC e controle de autoridade), o desempenho no trabalho (como analítico, comunicação, interpessoal, organizacional e solução de problemas), e a capacidade de trabalhar em um ambiente de equipe e de forma independente. Em termos de tecnologia da informação, eles devem entender *Extensible Markup Language* (XML), *Open Archives Initiative Protocol for Metadata Harvesting* (OAI-PMH) e *Resource Description Framework* (RDF), que são usados para conversão e compartilhamento de metadados. Ainda, os pesquisadores relatam que, a partir de 2004, os bibliotecários de metadados começaram a fazer parte de outros tipos de unidades ou departamentos dentro de uma biblioteca, como uma unidade de biblioteca digital, computação de biblioteca e serviços de mídia, e aquisição e gerenciamento de informações. A inserção de bibliotecários de metadados nesses espaços sugere que a posição de tal bibliotecário se originou em resposta às mudanças na forma como as bibliotecas estavam entregando e fornecendo acesso às suas coleções.

O contexto apresentado, que mostra a evolução do bibliotecário de catalogação em relação ao trabalho com metadados, exige do profissional familiaridade e experiência com uma variedade de formatos, padrões, esquemas,

ferramentas e práticas recomendadas. A posição do bibliotecário de metadados reflete a natureza em rápida mudança da catalogação na Biblioteconomia. Em dois artigos Gold (2007a e 2007b), já apontava como papel dos bibliotecários: criar metadados para descoberta e descrição de dados, organizando documentação para curadoria digital e oferecendo suporte de preservação de dados.

Kennan (2016), em estudo com bibliotecários atuando em universidades nos Estados Unidos, Reino Unido, Canadá e Austrália, listou os requisitos e habilidades indispensáveis para os bibliotecários de dados. O primeiro conjunto inclui habilidades interpessoais e características comportamentais, comunicação de alto nível (incluindo redação de documentação e outras redações técnicas). Outros atributos vinculados ao perfil profissional do bibliotecário de dados são a capacidade de estar continuamente atualizando-se, ter uma "curiosidade sem limites" e ser adaptável. O segundo conjunto de habilidades está relacionado ao conhecimento sobre o ambiente de pesquisa da universidade ou instituição de pesquisa, as políticas das agências financiadoras, as métricas da pesquisa e a avaliação. O terceiro conjunto refere-se a conhecimentos específicos sobre o uso de dados e o entendimento dos tipos de dados (quantitativo, qualitativo), licenças (*Copyright* e *Creative Commons*), preservação de dado e padrões e esquemas de metadados (*Dublin Core*, *Darwin Core*, *RDF*, *DIF*, *ANZLIC*, *RIF-CS*). A maioria dos participantes do estudo de Kennan (2016) enfatizou a necessidade desse conhecimento sobre a finalidade e o uso de metadados e

demonstrou uma vontade de aceitar e aprender os diferentes padrões e esquemas aplicáveis em diferentes contextos.

A gestão de dados de pesquisa abrange uma ampla gama de atividades durante todo o ciclo de vida dos dados. Geralmente, requer um alto nível de interação com os pesquisadores e, também, oferecer serviços de suporte, inclusive técnicos. O bibliotecário de metadados pode trabalhar na alfabetização em dados (padrões de metadados, formatos, etc.), auxiliar a identificar padrões e formatos de dados e padrões de metadados, gerenciar a coleta de dados e a descrição, estudar as melhores práticas para estruturas de dados (tipos, formatos, vocabulários, ontologias e metadados), auxiliar a encontrar informações sobre estruturas de dados, tipos, formatos, vocabulários, ontologias e metadados (Schmidt; Shearer, 2016).

Existem vários modelos de ciclo de vida de dados, como por exemplo o desenvolvido pelo Arquivo de Dados do Reino Unido, o *UK Data Archive Data Lifecycle*, no qual é dada ênfase ao armazenamento de dados bem organizados, bem documentados, preservados e compartilhados para a promoção da investigação científica (UK DATA ARCHIVE, 2012) ou o modelo do Data *Documentation Initiative* (DDI) que desenvolveu o *DDI Combined Lifecycle Model*, um modelo de ciclo de vida para dados de pesquisa, particularmente da área das Ciências Sociais ou ainda, o modelo do *Digital Curation Centre* (DCC), o *Curation Lifecycle Model* é um modelo mais complexo dentre os outros modelos de curadoria digital, preocupando-se em promover uma curadoria e preservação de dados bem-sucedida, a qual tem início desde a

conceituação dos dados até o descarte ou a seleção dos mesmos para reutilização e preservação a longo prazo.

Porém, acreditamos que o modelo de ciclo de vida dos dados elaborado pela *Data Observation Network of Earth* (DataONE), por ter foco mais específico em uma área, os dados da terra, seus oito componentes contemplam mais ampla e detalhadamente as questões relativas à importância dos metadados. A Figura 2 mostra os componentes do *DataONE Data Life Cycle*: planejar, coletar, assegurar, descrever, preservar, descobrir, integrar e analisar.

Figura 2: Ciclo de vida dos dados



Fonte: DataONE, [201-?].

De acordo com as informações da DataONE, existem várias versões de um ciclo de vida de dados com diferenças que variam conforme as práticas entre domínios ou comunidades (DATAONE, [201-?]). Porém, entre os componentes podemos identificar a importância da descrição dos metadados:

- a) planejar o arquivamento dos dados e a descrição dos mesmos: devem ser bem organizados, gerenciados e documentados para garantir que os arquivos possam ser localizados e seu conteúdo seja compreensível. Nesta etapa, é crucial identificar os dados sensíveis;
- b) coletar e prever como os dados serão descritos: os campos de dados devem ser separados de forma consistente por delimitadores aceitáveis (por exemplo, vírgulas) e usar esquemas de codificação consistentes, de forma a facilitar sua preservação e posterior acesso;
- c) assegurar a qualidade dos dados e os métodos de controle devem fazer parte do arquivo de informações de dados, preparar e documentar o plano de gestão de dados para garantir a confiabilidade e precisão. O controle de qualidade é altamente dependente da metodologia de coleta de dados, por esse motivo ela deve ser documentada com clareza;
- d) descrever os dados de forma minuciosa, por meio da adoção de padrões e ferramentas de metadados, além de permitir a interoperabilidade por meio dos metadados, possibilita que outros descubram, compreendam e usem esses dados. É necessário definir um dicionário de dados, no início do projeto, para descrever e documentar os dados com eficiência, por meio de metadados, que descrevem o contexto digital do conjunto de dados, o formato para localização geoespacial, formatos de data e

hora, as etapas de processamento, o conteúdo de um arquivo de dados com as informações sobre a equipe de coleta, o responsável pelo conjunto de dados, os patrocinadores da pesquisa. Ainda, para descrever o contexto científico (como foram coletados, quais os instrumentos usados), fornecer informações sobre os parâmetros (unidade de medida, formato do conjunto de dados, definição dos códigos usados), bem como, informar como o conjunto de dados deve ser citado. As taxonômicas completas para cada espécie devem ser incluídas na descrição dos dados;

- e) preservar é uma etapa do trabalho que é realizada em conjunto com os gestores do repositório de dados, os quais fornecem orientações sobre como gerar os metadados e preservar os dados e quais formatos usar. Não se pode esquecer a necessidade do uso de terminologias padrões (como palavras-chave, ontologias). Nesta etapa, também é necessário identificar a sensibilidade dos dados e fazer a classificação de segurança apropriada;
- f) descobrir configura-se no compartilhamento e publicação dos dados usando ferramentas e serviços baseados na *Web* para que as pessoas possam encontrar os dados e entender o seu conteúdo, assim como identificar outros conjuntos de dados e repositórios que possam complementar e agregar valor à determinada pesquisa. Para facilitar a descoberta, o nome dos arquivos deve ser descritivo e refletir o seu conteúdo;

- g) integrar os dados, a fim de que sejam compatíveis e comparáveis e documentados por meio de metadados. Todas as etapas realizadas para criar o conjunto de dados de uma pesquisa devem ser documentadas de forma completa e precisa para permitir que outros avaliem e interpretem seus resultados. Por exemplo, descrever os parâmetros geoespaciais (resolução, projeção, escala) de dados referenciados espacialmente ao integrar conjuntos de dados geoespaciais de fontes múltiplas;
- h) analisar consiste em integrar dados para que sejam compatíveis e comparáveis com outros. A descrição das etapas de processamento é imprescindível para converter uma observação em um produto de dados derivado. Devem ser documentadas de forma completa e precisa para permitir que outros avaliem e interpretem seus resultados.

Finalmente, cumpre-se o objetivo principal da descrição completa, precisa e consistente de metadados do conjunto de dados de uma pesquisa que, ao serem publicados, utilizando ferramentas e serviços como os repositórios de dados, possibilitam a reprodutibilidade, além de serem descobertos e acessados por qualquer pessoa. Organizar e descrever com metadados apropriados, armazenados em local onde possam ser monitorados e disponibilizados para uso futuro e preparados para migração ou transformação, caso o formato de dados esteja danificado ou obsoleto (Almeida, 2019), também se tornou uma das atividades primordiais dos profissionais da informação ou bibliotecários de metadados.

Mas, por outro lado, “[...] exige o desenvolvimento, a coordenação e investimentos em vários setores da biblioteca para criar um sistema onde, certamente, estarão envolvidos a segurança dos dados, a preservação, o acesso e o controle dos metadados.” (Cunha, 2010, p. 193).

5 CONSIDERAÇÕES FINAIS

Como um dos elementos que compõem o ecossistema da Ciência Aberta, os repositórios tornam-se fontes de busca de dados para pesquisas adicionais e cruzamento de dados, com vários níveis de complexidade para todos os atores envolvidos, porque não são estáticos como uma publicação científica. Apresentam desafios relacionados ao versionamento de arquivos e *softwares* para leitura e processamento, capacidade computacional e interação máquina-máquina, custos para armazenamento de *big data* e manutenção dos repositórios, perda de dados, preparo de *datasets* com a presença de dados pessoais, reusabilidade efetiva dos dados, duplicação de *datasets* em repositórios, indexação dos milhares de repositórios por serviços/mecanismos de busca (por exemplo, *Google Dataset Search*, *Data Citation Index*), certificação internacional, entre outros.

Especificamente para o profissional da informação, que desenvolve atividades focadas tanto no produtor de dados (seleção e avaliação de esquemas de metadados, capacitação e apoio técnico, preenchimento qualificado dos metadados) quanto nos usuários secundários (descoberta fácil dos dados, funcionalidades de busca e navegação do repositório, testes de

usabilidade), consideramos que entre os principais desafios estão a sobreposição de esquemas de metadados e o atendimento de necessidades específicas, de acordo com a área do conhecimento, especialmente em repositórios genéricos, sem perder de vista os princípios FAIR.

Na avaliação dos metadados, cuja atividade entendemos ser da competência do profissional da informação, deve-se considerar a quantidade e qualidade dos mesmos e sua adequação para descoberta de dados. Porém, Löffler *et al.* (2021) salientam que solicitar aos pesquisadores que forneçam “metadados ricos” para descrever seus conjuntos de dados, sem qualquer orientação, não resultará em algo satisfatório, pois muitas vezes não é claro quais são as informações relevantes, para a descoberta dos dados, que devem incluir no momento de submissão dos mesmos no repositório. Os autores recomendam documentar as práticas, principalmente aquelas aplicadas a domínios específicos, fornecer mais informações sobre os vários padrões de metadados e as categorias de pesquisa por campo de busca.

Segundo Almeida (2019), às bibliotecas cabem agora, dentre outros serviços, a definição de padrões de metadados e a realização de treinamentos voltados para a gestão de dados oriundos de pesquisas. Dessa forma, é importante que o profissional da informação esteja próximo aos produtores de dados e engajado ativamente nas atividades de curadoria dos dados e de capacitação e suporte aos pesquisadores, com a finalidade de compartilhar dados de pesquisa em acesso aberto, atendendo aos princípios FAIR, e de modo a qualificar os metadados e, conseqüentemente, os próprios dados. O uso

e reuso dos dados de pesquisa, de forma adequada, estão diretamente relacionados com a qualidade com que são descritos.

REFERÊNCIAS

ALMEIDA, F. G. *Suporte à gestão de dados de pesquisa: uma ampliação dos serviços oferecidos pelas bibliotecas*. 2019. Tese (Doutorado em Gestão & Organização do Conhecimento) - Escola de Ciência da Informação, Universidade Federal de Minas Gerais, Belo Horizonte, 2019. Disponível em: <http://hdl.handle.net/1843/31596>. Acesso em: 19 jul. 2021.

ÁVILA BARRIENTOS, E. Análisis de los principios para la descripción de datos de investigación mediante datacite metadata schema. *Anales de documentación*, Múrcia, v. 23, n. 1, 2020. Disponível em: <http://dx.doi.org/10.6018/analesdoc.400341>. Acesso em: 29 jun. 2021.

BOERES, S. A. de A. *Competências necessárias para equipes de profissionais de preservação digital*. 2017. Tese (Doutorado em Ciência da Informação) - Faculdade de Ciência da Informação, Universidade de Brasília, Brasília, 2017. Disponível em: <https://repositorio.unb.br/handle/10482/24354>. Acesso em: 19 jul. 2021.

BORGMAN, C. L. *Big data, little data, no data: scholarship in the networked*. Cambridge: MIT Press, 2015.

CAMPBELL, G. D. Metadata, Metaphor, and Metonymy. In: SMIRAGLIA, Richard. *Metadata: A Cataloger's Primer*. New York: Haworth Information, 2006. p. 4-5.

CHAPMAN, A. *et al.* Dataset search: a survey. *The VLDB Journal*, Heidelberg, v. 29, p. 251-272, 2020. Disponível em: <https://arxiv.org/abs/1901.00735>. Acesso em: 19 jul. 2021.

CORRÊA, F. C. *Gestión de datos de investigación*. Barcelona: Editorial UOC, 2016.

CUNHA, M. B. da. A biblioteca universitária na encruzilhada. *DataGramZero*: revista de ciência da informação, Rio de Janeiro, v. 11, n. 6, dez. 2010. Disponível em: <https://repositorio.unb.br/handle/10482/14869>. Acesso em: 19 jul. 2021.

DATACITE METADATA WORKING GROUP. *DataCite metadata schema documentation for the publication and citation of research data and other research outputs*. Version 4.4. DataCite e.V. 2021. Disponível em: <https://doi.org/10.14454/3w3z-sa82>. Acesso em: 15 jun. 2021.

DATAONE. *Data life cycle*. [201-?]. Disponível em: <https://old.dataone.org/data-life-cycle>. Acesso em: 19 jul. 2021.

GOLD, A. Cyberinfrastructure, Data, and Libraries, Part 1; A Cyberinfrastructure Primer for Librarians. *D-Lib Magazine*, Reston, v. 13 n. 9/10, 2007a. Disponível em: <http://www.dlib.org/dlib/september07/gold/09gold-pt1.html>. Acesso em: 19 jul. 2021.

GOLD, A. Cyberinfrastructure, data, and libraries, part 2: Libraries and the data challenge: Roles and actions for libraries. *D-Lib Magazine*, Reston, v. 13 n. 9/10, 2007b. Disponível em: <http://www.dlib.org/dlib/september07/gold/09gold-pt2.html>. Acesso em: 19 jul. 2021.

HAN, M-J.; HSWE, P. The evolving role of the metadata librarian: competencies found in job descriptions. *Library resources & technical services*, Chicago, v. 54, n. 3, p. 129-141, July 2010. Disponível em: <https://journals.ala.org/index.php/lrts/article/view/5557/6839>. Acesso em: 19 jul. 2021.

KENNAN, M. A. Data management: knowledge and skills required in research, scientific and technical organisations. *In: IFLA WORLD LIBRARY AND INFORMATION CONGRESS, 2016. Columbus. Proceedings...* [S. l.]: IFLA, 2016. Disponível em: <http://library.ifla.org/1466/1/221-kennan-en.pdf>. Acesso em: 19 jul. 2021.

KHAN, S. M. *et al.* A global review of publicly available datasets for ophthalmological imaging: barriers to access, usability, and generalisability. *Lancet digital health*, Oxford, v. 3, n. 1, e51–66, jan., 2021. Disponível em: [https://doi.org/10.1016/S2589-7500\(20\)30240-5](https://doi.org/10.1016/S2589-7500(20)30240-5). Acesso em: 05 jul. 2021.

KOESTEN, L. M. *et al.* The trials and tribulations of working with structured data - a study on information seeking behaviour. *In: CHI CONFERENCE ON HUMAN FACTORS IN COMPUTING SYSTEMS, 2017, Denver. Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. New York: ACM, 2017. p. 1277-1289. Disponível em: <https://doi.org/10.1145/3025453.3025838>. Acesso em: 03 jul. 2021.

LÖFFLER, F. *et al.* Dataset search in biodiversity research: Do metadata in data repositories reflect scholarly information needs? *PLoS ONE*, San Francisco, v. 16, n. 3, e0246099, 2021. Disponível em: <https://doi.org/10.1371/journal.pone.0246099>. Acesso em: 03 jul. 2021.

MONTEIRO, E. C. de S. de A.; SANT'ANA, R. C. G. Plano de gerenciamento de dados em repositórios de dados de universidades. *Encontros Bibli: revista eletrônica de biblioteconomia e ciência da informação*, Florianópolis, v. 23, n. 53, p. 160-173, set./dez., 2018. Disponível em: <https://doi.org/10.5007/1518-2924.2018v23n53p160>. Acesso em: 28 jun. 2021

NATIONAL INFORMATION STANDARDS ORGANIZATION. *A Framework of guidance for building good digital collections*. 3rd. ed. Baltimore: NISO, 2007. Disponível em: <http://www.niso.org/publications/rp/framework3.pdf>. Acesso em: 05 abr. 2021.

ORGANIZATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT. *OECD principles and guidelines for access to research data from public funding*. Paris: OECD, 2007. Disponível em: <https://www.oecd.org/sti/inno/38500813.pdf>. Acesso em: 24 jun. 2021.

PAVÃO, C. M. G. *et al.* Metadados e repositórios institucionais: uma relação indissociável para a qualidade da recuperação e visibilidade da informação. *PontodeAcesso*, Salvador, v. 9, n. 2, p. 103-116, 2015. Disponível em: <http://hdl.handle.net/10183/175147>. Acesso em: 25 jun. 2021.

PONTIKA, N. *et al.* Fostering open science to research using a taxonomy and an eLearning portal. *In: INTERNATIONAL CONFERENCE ON KNOWLEDGE TECHNOLOGIES AND DATA-DRIVEN BUSINESS*, 15., 2015, Graz. *Proceedings [...]*. New York: ACM, 2015. Article n. 11. Disponível em: <https://doi.org/10.1145/2809563.2809571>. Acesso em: 28 jun. 2021.

ROCHA, R. P. da *et al.* *Acesso aberto a dados de pesquisa no Brasil: planejamento para implantação de comunidade*

produtora de dados para o repositório Rede de Dados de Pesquisa. 2019. Disponível em: <http://hdl.handle.net/20.500.11959/1204>. Acesso em: 25 jun. 2021.

ROUSIDIS, D. *et al.* Metadata for big data: a preliminary investigation of metadata quality issues in research data repositories. *Information services and use*, Amsterdam, v. 34, n. 3-4, p. 279-286, 2014. Disponível em: <https://content.iospress.com/download/information-services-and-use/isu746?id=information-services-and-use%2Fisu746>. Acesso em: 19 jun. 2021.

ROUSIDIS, D. *et al.* Evaluation of metadata in research data repositories: the case of the DC.Subject element. *In*: GAROUFALLOU, E.; HARTLEY, R.; GAITANOU, P. (eds). *Metadata and semantics research*. [S. l.]: Springer, 2015. Disponível em: https://doi.org/10.1007/978-3-319-24129-6_18. Acesso em: 19 jun. 2021.

SALES, L. F.; SAYÃO, L. F. *Dados de pesquisa: quem ama cuida*. Brasília: Comissão Nacional de Energia Nuclear: IBICT, 2019. Disponível em: <http://livroaberto.ibict.br/handle/123456789/1083>. Acesso em: 25 jun. 2021.

SCHMIDT, B.; SHEARER, K. *Librarians' Competencies Profile for Research Data Management*. Joint Task Force on Librarians' Competencies in Support of EResearch and Scholarly Communication. 2016. Disponível em: https://www.coar-repositories.org/files/Competencies-for-RDM_June-2016.pdf. Acesso em: 19 jul.2021.

SOLODOVNIK, I. Development of a metadata schema describing institutional repository content objects enhanced by "LODE-BD" strategies. *JLIS.it.*, Macerata, v. 4, n. 2, p. 109-

144, 2013. Disponível em: <http://dx.doi.org/10.4403/jlis.it-8792>. Acesso em: 28 jun. 2021.

UNESCO. *Hacia una recomendación de la UNESCO sobre la ciencia abierta: crear un consenso mundial sobre la ciencia abierta*. Paris: UNESCO, 2020. Disponível em: https://unesdoc.unesco.org/ark:/48223/pf0000373209_spa. Acesso em: 24 jun. 2021.

WHITE, H. C. Descriptive metadata for scientific data repositories: a comparison of information scientist and scientist organizing behaviors. *Journal of library metadata*, New York, v. 14, n. 1, p. 24-51, 2014. Disponível em: <https://doi.org/10.1080/19386389.2014.891896>. Acesso em: 19 jun. 2021.

WILKINSON, M. D. *et al.* The FAIR Guiding Principles for scientific data management and stewardship. *Scientific data*, Londres, v. 3, article 160018, 2016. Disponível em: <https://doi.org/10.1038/sdata.2016.18>. Acesso em: 02 jul. 2021.

WU, M. *et al.* Data discovery paradigms: user requirements and recommendations for data repositories. *Data science journal*, London, v. 18, article 3, 2019. Disponível em: <https://doi.org/10.5334/dsj-2019-003>. Acesso em: 19 jun. 2021.

CAPÍTULO 3

CATALOGAÇÃO DE ASSUNTO: REFLEXÕES DE TERMINOLOGIA DESDE A DÉCADA DE 1960

Rosali Fernandez de Souza

Naira Christofoletti Silveira

1 INTRODUÇÃO

O II Encontro de Representação Documental (EnReDo) tem como tema a Representação Documental e o ambiente Web, com enfoque para os desafios e oportunidades da Organização e Representação da Informação na era digital. Aproveitamos esta ocasião para refletir sobre Eric Coates, bibliotecário e classificador inglês, que teve como grande desafio nas décadas de 1960 e 1970 o início da mecanização do *British Technology Index*.

A obra de Eric Coates que marca este momento se intitula *Subject Catalogues: Headings and Structure* e foi publicada pela primeira vez, em Londres, em 1960 e reimpressa em 1963 e 1969. Na introdução do livro, Coates (1969) apresenta a seção *Terminology* que compreende 21 termos que permeiam o livro como um todo e 11 termos que são evidentes por si, mas que estão relacionados para apresentar ao leitor um referencial em caso de interpretação equivocada. Neste sentido, o autor apresenta 32 termos que julgou relevantes, relacionados diretamente ao início do uso dos computadores na denominada catalogação de assunto.

Considerando os termos trazidos por Coates (1969), buscou-se a aderência de tais definições na *Statement of International Cataloguing Principles* (ICP) e nos *Functional Requirements for Subject Authority Data* (FRSAD), documentos publicados pela *International Federation of Library Associations and Institutions* (IFLA).

O nosso objetivo é retomar a terminologia da catalogação de assunto indicada por Coates no início da década de 1960, e discuti-la à luz da atualidade, após mais de 50 anos. A partir da análise dos termos é possível acompanhar o desenvolvimento da catalogação de assunto enquanto atividade indispensável, que sempre foi, continua sendo e sempre será, no tratamento de dados, informação e documentos em diferentes tipos de sistemas e serviços de informação em áreas, campos e domínios do conhecimento.

Em um primeiro momento, apresentamos a trajetória de Eric Coates; depois os conceitos relacionados à catalogação de assunto destacados por ele, partindo de uma análise crítica sobre os termos, conceitos e definições de tais termos na atualidade. Por fim, nossas considerações finais.

2 COATES E A CATALOGAÇÃO DE ASSUNTO

Para Guimarães (2008, p. 82), o Tratamento Temático da Informação (TTI) ocorre a partir de três abordagens: “[...] *subject cataloguing* (de orientação predominantemente norte-americana), *indexing* (de orientação predominantemente inglesa) e *analyse documentaire* (de orientação predominantemente francesa).” Embora de orientação

predominante norte-americana, a catalogação de assunto, conforme autor acima citado, tem significativa contribuição dos trabalhos de Cutter, Kaiser e Coates.

Sendo assim, neste texto, adotamos o termo catalogação de assunto, seguindo a definição de Coates (1969), que além de pesquisador deixa um legado importante e de destaque na indexação, pois atuou aliando a teoria à prática. Portanto, resgatar a terminologia adotada por Coates no início da automação contribui para compreender a Catalogação no âmbito da Organização e Representação de documentos na atualidade.

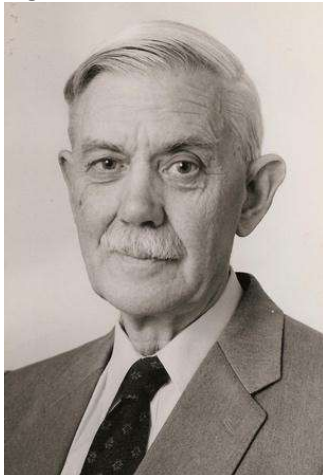
2.1 Eric Coates e a sua trajetória

Eric James Coates, nasceu em 1916 e faleceu em 2017. Foi membro fundador do *Classification Research Group* (CRG) e entre as suas principais linhas de atuação estão a aplicação da teoria na prática de construção e uso de linguagens de indexação (Kawamura, 2020).

Influenciado pelo pensamento de Ranganathan, foi um dos pioneiros a adotar o método de análise de facetas fora da Índia. Ele também colaborou com a equipe que desenvolveu o *Broad System of Ordering* (BSO) construído pela *Fédération Internationale de Documentation* (FID), em associação com a UNESCO no âmbito do programa UNISIST, que pretendia ser um mecanismo de mudança para várias linguagens de indexação. A expertise de Coates fez com que a versão preliminar do BSO fosse refinada, concluída e submetida a teste de aplicação. Com base nas descobertas e na experiência desses testes de campo, as revisões posteriores do BSO se

desenvolveram. A teoria e a prática de Coates tiveram sempre como base a seguinte tríade de princípios: a unidade básica de indicação de assunto; análise relacional no contexto da classificação; e, recorrer ao suporte da classificação em qualquer caso (Kawamura, 2020).

Figura 1: Eric James Coates



Fonte: Kawamura (2020).

Na sua trajetória profissional e de pesquisador, Coates se deparou com dois importantes desafios: a computadorização do *British Technology Index* (BTI), periódico em que atuava como editor e o seu envolvimento num projeto internacional para desenvolver um léxico intermediário para permitir a fácil interconversão entre listas de vocabulário controlado, tanto em ordem alfabética quanto classificada, usadas para indexar literatura no campo da Documentação e Biblioteconomia. O projeto foi denominado “*Documentation of documentation*”, iniciado pelo *Groupe d’ Etude sur l’Information Scientifique* (GEIS) de Marselha, chefiado por Jean-Claude Gardin (Kawamura, 2020).

Em seus principais artigos publicados no final da década de 1980, Coates começou a usar a expressão 'linguagem de informação', ao invés de linguagem de indexação, mas o significado da mudança na terminologia não foi adotado. Coates usou essa expressão para significar as seguintes proposições:

- as linguagens de informação são a base de todos os métodos de recuperação;
- a classificação é ou deve ser a base para as linguagens de informação (Kawamura, 2020).

Com base nessas duas proposições, Coates evidenciou dois desempenhos adicionais de classificação: um auxílio de pesquisa na recuperação e um auxílio de ensino na educação em geral. A primeira ideia baseia-se no pressuposto de que a recuperação online pelos próprios usuários será ainda mais popularizada e que uma linguagem de indexação se tornará uma ferramenta útil para a recuperação. A segunda ideia remonta ao discurso feito por Ranganathan na *Dorking Conference*, em 1957, quando argumentou que as universidades deveriam manter cursos e bolsas de pesquisa na disciplina de classificação de bibliotecas (Kawamura, 2020).

Widad Mustafa El Hadi (2015), baseada em Coates, considera que combinando os dois tipos de Sistema de Organização do Conhecimento (SOC) - classificações e linguagem, como tesouros e vocabulários multilíngues - poderia se garantir a interoperabilidade, reafirmando, assim, que além dos esquemas de classificação deveriam existir outros tipos de linguagens.

Kawamura (2004) aponta três critérios propostos por Coates para o *British Technology Index* (BTI): cobertura, atualidade e poder de recuperação (*coverage, currency and retrieving power*). Observa-se que até hoje são critérios correntes no que se refere à indexação, ou seja:

A coerência conceitual subjacente no trabalho de Coates deve ser altamente considerada como a sobrevivência persistente de interesse e preocupação com a classificação, apesar de sua marginalização, tanto na comunidade de bibliotecas e ciências da informação e mais amplamente, pela gestão empresarial dominante, a cultura de tecnologia da informação dos dias atuais. [...] ainda um outro exemplo é o surgimento da classificação, sob nomes como ontologia e taxonomia, como uma preocupação viva dos usuários e provedores de serviços da Internet. [...] Classificar é a formação de classes por referência às propriedades de objetos [...]. Permeia o próprio processo de pensamento em todas as fases da vida. (Kawamura, 2004, p. 342, tradução nossa)

Enfim, podemos dizer que a prática de uma linguagem de indexação, aliada a uma educação em classificação, contribuirá para melhorar a efetividade de cada usuário na recuperação na era de uma sociedade de informação, onde a capacidade de identificar e navegar com termos na diversidade de campos de conhecimento seria uma habilidade necessária. Nesse sentido, no pensamento de Coates, novas classificações gerais completamente facetadas, baseadas numa visão de mundo moderna que também preenchem a lacuna entre a teoria e a prática da classificação, são os instrumentos mais

valiosos para alcançar os dois aspectos de classificação acima mencionados.

As atuações de Coates o levaram a publicar o já mencionado livro *Subject Catalogues: headings and structure*, como uma nova abordagem para catálogos alfabéticos de assunto. Na edição de 1969, o capítulo 1, é dedicado à Terminologia reconhecendo ser esta imprecisa no tema catálogo de assunto quanto a cabeçalhos e estrutura. Por esse motivo, nós nos sentimos motivadas a resgatar termos-chave listados nessa Terminologia que consideramos relevantes revisitar para reflexão no contexto atual da catalogação de assunto.

2.2 Termos relacionados à catalogação de assunto

Na seleção de termos que formou o capítulo “Terminologia” é interesse destacar o entendimento do raciocínio do autor no conjunto das duas sublistagens de termos que apresenta: uma em que são listados termos típicos e consolidados conceitualmente na literatura da época, e outra em que apresenta termos que podem induzir a confusões terminológicas e a interpretações equivocadas pelos leitores.

A listagem 1 corresponde ao conjunto dos termos considerados por Coates como consolidados na atividade da catalogação de assunto na década de 1960 e, como tal, permearam a obra em seu todo. São eles: *Catalogue and Index*, *Document*, *Enquirer*, *Subject catalogue*, *Alphabetic-specific subject catalogue*, *Alphabetic-classed subject catalogue*, *Classified catalogue*, *Subject entry*, *Subject heading*, *Natural language*, *Feature heading*, *Entry word*, *Subheading*,

Compound subject, Component, Reference, Upward reference, Downward reference, Collateral reference, Mechanical selector, Neutrality.

Os termos são evidentes por si na atividade de catalogação de assunto, especificamente relacionados aos cabeçalhos e à estrutura, a proposta do livro. A própria ordenação desses termos sugere uma sistemática de raciocínio do autor, com termos que revelam elementos tanto de natureza contextual geral, quanto de natureza específica e especializada envolvidos na catalogação de assunto. Parte da sinonímia entre **Catálogo e Índice** e segue trazendo os dois referenciais básicos de qualquer sistema de informação aos quais a catalogação de assunto é destinada, ou seja, o **Documento** e o **Usuário** (referido como inquiridor). Enumera tipos de catálogos: **Catálogo de Assunto, Catálogo alfabético de assunto específico, Catálogo de assunto classificado em ordem alfabética e Catálogo Classificado**. Em seguida, os termos específicos de construção de catálogos: **Entrada de Assunto e Cabeçalho de Assunto** e os principais elementos que envolvem essa construção: **Linguagem natural e Cabeçalho especial** definem a **Palavra de Entrada** e os **Subcabeçalhos**.

Seguindo o raciocínio de Coates, o autor passa então a sinalizar para questões, dificuldades e dúvidas, sobre a definição de assunto, ou seja, **Assunto Composto e Componente**. Adentra para aspectos tipicamente de estrutura quando elenca os termos que relacionam a assuntos correlatos, tais como: **Referente, Referente Ascendente, Referente Descendente e Referente Colateral**. Termina com aspecto de operabilidade física, o **Seletor Mecânico**, que se

evidenciava à época, e o desafiador aspecto de natureza intelectual - a **Neutralidade** - tanto para quem constrói o catálogo, o catalogador, quanto para quem usa catálogo de assunto: o próprio catalogador e o usuário do sistema de informação.

A listagem 2 apresenta o conjunto de termos que são de uso corrente com significação reconhecida. No entanto, foram considerados como referencial no contexto da classificação que envolve catalogação de assunto, evitando causar interpretação conceitual equivocada. São eles: *Schedule, Facet, Chain, Chain procedure, Qualifier, Basic analysis, Qualified list, Approach term, Hierarchical or Generic search - Substitution generic, Systematic colateral search, Alphabetical colateral search.*

O **Esquema** de classificação como o instrumento de apoio à catalogação de assunto. Os termos **Faceta**, **Cadeia** e **Procedimento em cadeia** e **Qualificador** representam métodos para embasar a construção de cabeçalho de assunto e, como abordagem temática, aponta **Análise básica** e **Lista qualificada**. Do ponto de vista da recuperação, enumera: **Termo de abordagem** e diferentes tipos de busca: **Busca hierárquica ou genérica - Substituição genérica, Busca colateral sistemática, Busca colateral alfabética.**

A seguir apresentamos breve definições de tais termos propostos por Coates (1969), em tradução livre:

1. **Catálogo e índice:** Catálogo - especificações descritivas de assunto; Índice - localização de assunto;
2. **Documento:** registro gráfico, acústico ou outro tipo;

3. **Usuário:** consulente do catálogo de assunto;
4. **Catálogo de assunto:** termos de assunto para um documento em ordem alfabética ou sistemática;
5. **Catálogo de assunto específico em ordem alfabética:** cabeçalho de assuntos organizados em ordem alfabética; a parte de assunto de um catálogo dicionário;
6. **Catálogo de assunto classificado em ordem alfabética:** entradas de assuntos específicos, subordinados em ordem alfabética;
7. **Catálogo classificado:** entradas de assuntos em ordem sistemática, exibindo relação hierárquica entre assuntos no mesmo nível em ordem de complexidade, cronológica, pedagógica ou alfabética; catálogo sistemático: catálogo classificado de entradas em ordem sistemática, e complementar índice de assunto alfabético;
8. **Entrada de assunto:** unidade básica do catálogo de assunto;
9. **Cabeçalho de assunto:** declaração do assunto de um documento (ou outra unidade literária);
10. **Linguagem natural:** circunstâncias particulares de cada idioma - necessitam de qualificação;
11. **Cabeçalho especial:** parte verbal do cabeçalho de assunto usado num arquivo sistemático de um catálogo classificado;
12. **Palavra de entrada:** palavra principal de um cabeçalho de assunto em um catálogo ou índice organizado em ordem alfabética;
13. **Subcabeçalho:** segunda palavra de assunto verbal, separada da palavra precedente por pontuação;

14. **Assunto composto:** assunto que requer mais de uma palavra para expressar o seu significado num cabeçalho;
15. **Componente:** palavra individual constituinte num cabeçalho de assunto composto;
16. **Referente:** direciona um cabeçalho para outro na mesma sequência;
17. **Referente ascendente:** direciona um cabeçalho de assunto de menor abrangência para um de maior abrangência;
18. **Referente descendente:** direciona um cabeçalho de maior abrangência para um de menor abrangência;
19. **Referente colateral:** liga dois cabeçalhos pertencentes ao mesmo nível hierárquico de um termo genérico comum;
20. **Seletor mecânico:** catálogo de assunto incorporando a mecanização das entradas para busca por assunto; seletores mecânicos definem código como princípio de busca para encontrar correspondente num armazenamento de entradas codificadas de assunto;
21. **Neutralidade:** preferência do usuário não pode ser invocada (seja porque é incerta ou porque não existe) para justificar um curso de ação sobre outro;
22. **Esquema:** lista de termos em relacionamento hierárquico (tabelas de um esquema de classificação);
23. **Faceta:** termos de assunto (ou conceitos de assunto) de característica única; facetas podem ser divididas em subfacetas;
24. **Cadeia:** hierarquia de termos num esquema de classificação, cada termo contendo ou incluindo todos aqueles que o seguem;

25. **Procedimento em cadeia:** método de construir entradas de índice de assuntos, sem permutação de componentes, citando termos contidos em cadeias particulares;
26. **Qualificador:** subcabeçalho num índice de assunto, construído por procedimento em cadeia;
27. **Análise básica:** primeiro estágio na conversão de um símbolo de classificação num cabeçalho de assunto verbal para um catálogo alfabético específico; citação em ordem hierárquica ascendente dos elementos constituintes de um assunto composto;
28. **Lista qualificada:** expressão usada para denotar o segundo estágio da conversão de um símbolo de classe num cabeçalho de assunto verbal; é produzida pela exclusão de elos Falsos ou Não Procurados da cadeia e adicionando qualificadores aos termos restantes;
29. **Termo de abordagem:** única palavra que o usuário busca no catálogo que o levará a uma declaração de assunto composto;
30. **Busca hierárquica ou genérica:** análise das entradas sob os cabeçalhos que constituem uma cadeia; conduzida na direção ascendente a partir do cabeçalho mais específico para o menos específico; o curso de uma pesquisa hierárquica pode incluir o que é denominado uma Substituição genérica, que denota uma situação no catálogo alfabético-específico em que dois cabeçalhos, cada um com o mesmo número de subcabeçalhos, representam conceitos em relação hierárquica;
31. **Busca colateral sistemática:** análise das entradas sob os cabeçalhos ficam ao lado de cada um dos cabeçalhos cobertos na busca hierárquica;

32. Busca colateral alfabética: análise das entradas sob os cabeçalhos adjacentes em ordem alfabética, ou nas proximidades dos cabeçalhos cobertos na busca hierárquica.

Embora muito interessante, não seria possível aprofundar as definições de todos os termos neste trabalho, por isso realizou-se um recorte de termos adotados por Coates que estejam presentes ou que permeiam os documentos publicados pela *International Federation of Library Associations and Institutions* (IFLA), relacionados a catalogação de assunto, a saber: *Statement of International Cataloguing Principles* (ICP) e *Functional Requirements for Subject Authority Data* (FRSAD).

3 OS TERMOS NA CATALOGAÇÃO DE ASSUNTO

A primeira análise se refere aos termos que permeiam o livro de Coates (1969) como um todo e que constam na *Statement of International Cataloguing Principles* (ICP), publicado pela IFLA (2016).

No glossário da ICP apenas o termo “usuário” foi identificado. Ele tem o sentido que Coates apresenta como “*Enquirer*”. O termo “catálogo de assunto” não consta, e o termo Catalogação é abordado de modo amplo, “catalogação descritiva” e “catalogação por assunto”, em seu glossário com as seguintes definições:

Catalogação descritiva [Descriptive cataloguing]
– A parte da catalogação que proporciona informação descritiva e pontos de acesso, excluindo a dos assuntos. [...].

Catálogo por assunto [Subject cataloguing] –
 A parte da catalogação que identifica o *thema* e
 refere a ele ou *nomen*. (IFLA, 2016, p. 14)

Embora a ICP aborde a Catalogação como um todo, existe um documento que aborda especificamente a catalogação de assunto, os FRSAD.

A seguir são apresentados os termos analisados neste trabalho, considerados equivalentes ou que foram apresentados por Coates, presentes na ICP e no FRSAD, diretamente impactados com o uso dos computadores na catalogação de assunto, conforme quadro abaixo:

Quadro 1: Termos relacionados à catalogação de assunto.

TERMOS A SEREM ANALISADOS		
Coates	ICP	FRSAD
<i>Document</i>	<i>Resource</i>	<i>Resource</i>
<i>Enquirer</i>	<i>User</i>	<i>User</i>
<i>Subject Heading</i>	<i>Authorized access point, Controlled access point</i>	<i>Subject heading, Controlled access point</i>
<i>Reference</i>	<i>Variant form of name</i>	<i>Reference, Relationship, Hierarchical relationships</i>
<i>Upward reference</i>	<i>Relationships</i>	
<i>Downward reference</i>		
<i>Collateral reference</i>		

Fonte: Coates (1969) e IFLA (2010, 2016).

Considerou-se para esta análise os termos em comum ou que mostram de alguma forma uma trajetória que impacta os conceitos adotados hoje em catalogação de assunto. Os termos que permitem traçar um desenvolvimento conceitual ou mudança significativa para a catalogação de assunto de modo explícito são: *Document*, *Enquirer*, *Subject Heading*,

Reference, Upward reference, Downward reference, Collateral reference.

- a. **Document.** Observa-se um desuso do termo “documento” substituindo-o pelo termo “recurso”. Na ICP considera-se um recurso bibliográfico qualquer entidade, tangível ou intangível, que coleta conteúdo intelectual e/ou artístico (IFLA, 2016). Neste sentido, o termo documento poderia ter sido mantido, porém a adoção do termo “recurso” parece mais adequada à realidade computacional. Este termo inclusive está no substituto do Código de Catalogação Anglo-Americano (AACR2r), que se denomina *Resource Description and Access* (RDA);
- b. **Enquirer.** Coates ao longo de sua obra também adota o termo cliente e usuário. Usuário é o termo mais usual agora, definido pela ICP como qualquer pessoa, entidade coletiva ou dispositivo tecnológico que pesquisa no catálogo e faz uso dos dados bibliográficos e/ou de autoridades (IFLA, 2016). Interessante observar que a ICP coloca o usuário como centro do catálogo, como se ele estivesse apenas surgido recentemente na Catalogação, porém Coates na década de 1970 já evidenciava a importância de se definir o usuário do sistema;
- c. **Subject Heading.** A ICP não adota mais o termo *heading*, mas sim *Authorized access point* e *Controlled access point*. O FRSAD adota *subject heading* e *controlled access point*. O termo linguagem natural, por mais que não seja adotado formalmente, estaria

relacionado às formas não controladas de ponto de acesso, sendo: “Ponto de acesso autorizado”, o ponto de acesso controlado preferido para uma entidade, estabelecido e construído de acordo com as regras ou normas; “Ponto de acesso controlado”, o ponto de acesso indicado em um registro de autoridade; “Ponto de acesso não controlado”, o ponto de acesso que não está controlado por dados de autoridade (IFLA, 2016). Em relação ao “cabeçalho de assunto” observa-se muitos desdobramentos: palavras-chave, termo, descritor; esse termo por si só traz uma bagagem muito relevante para a construção de tesouros e ontologias, amplamente adotadas no meio digital;

- d. **Reference.** O termo “referente”, “referência” ou “remissiva” não é mais adotado atualmente, sendo seu substituto “Forma variante do nome”, que seria “Uma forma de nome da entidade não estabelecido como ponto de acesso autorizado e que se pode utilizar para acessar os dados de autoridade da entidade ou que se pode apresentar como um vínculo com o ponto de acesso autorizado” (IFLA, 2016, p. 15). Os vínculos presentes em um catálogo também podem ser entendidos como relacionamentos, que seriam “Uma conexão específica entre entidades ou suas exemplificações” (IFLA, 2016, p. 18). Os relacionamentos também estão previstos no FRISAD (2010), em especial quanto à questão hierárquica, que são estruturas que mostram relacionamentos entre conceitos e classes de conceitos e revelam graus ou níveis de superordenação e subordinação, em que o

termo superordenado representa uma classe ou um todo, e os termos subordinados se referem a seus membros ou partes.

Ao acompanhar os termos adotados em diferentes momentos, pode-se identificar o conjunto dos sistemas da organização do conhecimento. Por exemplo, o termo “cabeçalho de assunto”, para “descriptor”; é possível identificar a história e construção do desenvolvimento da própria catalogação de assunto.

Das listagens 1 e 2, como apoio a desenvolvimentos futuros da catalogação de assunto, destacamos como relevantes o termo Seletor mecânico e, os termos Referente e os respectivos tipos de Referentes. Em relação ao termo Seletor mecânico destacamos a mensagem de alerta de Coates já desde a década de 1960:

a recuperação de informação por máquina é ela própria confrontada por problemas de estrutura lógica, cuja invenção se tornou mais urgente. A máquina, não menos que sua contraparte manual, precisa de um programa lógico para ser utilizada com eficácia. (Coates, 1969, p. 175, tradução nossa)

Em relação aos termos Referente e seus correlatos: ascendente, descendente e colateral, embora Coates, não adote a palavra tesouros, fica indicada a atuação dos procedimentos de construção de tesouros como auxílio relevante para a catalogação de assunto.

4 CONSIDERAÇÕES FINAIS

Considerando os últimos 50 anos de início da computação em bibliotecas, especificamente das décadas de 1960 e 1970 até os anos 2021, muitas mudanças tecnológicas impactaram a catalogação de assunto, assim como os sistemas de organização do conhecimento (SOC) que se diversificaram e se tornaram mais complexos.

A mecanização se tornou cada vez mais sofisticada em termos das tecnologias de informação disponíveis, o que aumentou consideravelmente a rapidez de busca, exigindo recursos de interoperabilidade entre bases de dados e interfaces amigáveis com o usuário.

Por outro lado, novas abordagens de definição de assunto, principalmente quanto a conceitos, vêm revolucionando as formas classificatórias empregadas nos diferentes tipos de SOC, tais como: listas de cabeçalhos de assunto, vocabulários controlados, esquemas de classificação, tesouros e ontologias, para citar alguns.

Nesse sentido, os 32 termos elencados por Coates e as respectivas definições, embora ainda válidas em sua essência, tiveram sua terminologia adaptada no tempo e as definições se tornaram mais sofisticadas, em função da adequação aos novos contextos computacionais e de adaptações e revisões no desenvolvimento de SOC. Pelo seu legado de contribuições, podemos afirmar que Coates é uma das mais proeminentes figuras na teoria e na prática em classificação, indexação e recuperação de informação e que reverbera até hoje em sistemas automatizados.

Poderíamos acrescentar outros termos, ampliando, mantendo, modificando ou formando uma terminologia mais adequada à era digital. Mas o fato é que a preocupação apresentada por Coates sobre a terminologia presente em sua obra *Subject Catalogues: headings and structure* ainda está presente nos dias atuais. Observa-se uma certa fragilidade conceitual, ou mesmo uma tendência, em se adotar termos novos distanciados de teoria, valorizando ações práticas para delimitar conceitos.

A inquietação provocada pela pesquisa está na capacidade de identificar até que ponto a mudança terminológica representa uma necessidade conceitual e até que ponto tais alterações terminológicas podem camuflar ou provocar um certo apagamento de teóricos da Documentação e da Biblioteconomia, em prol de conceitos mais tecnológicos. Neste caso, destaca-se o desuso do termo “documento” por “recurso” e “inquiridor” para “usuário”.

Ao analisar a terminologia da catalogação de assunto indicada por Coates no início da década de 1960, e discuti-la à luz da atualidade após mais de 50 anos, fica claro que o sistema de informação evolui para rever e buscar minimizar o problema identificado no sistema anterior. Algumas limitações operacionais foram solucionadas com o uso de tecnologia, outras permanecem, em especial as questões conceituais que revelam uma preocupação terminológica em torno da catalogação de assunto e um vasto campo ainda a ser estudado.

REFERÊNCIAS

- COATES, E. J. *Subject Catalogues: Headings and Structure*. Reprinted. London: Library Association, 1969.
- EL HADI, W. M. Cultural Interoperability and Knowledge Organization Systems. In: GUIMARÃES, José Augusto; DODEBEI, Vera (org.). *Organização do Conhecimento e diversidade Cultural*. São Paulo: ISKO Brasil, 2015. (Estudos Avançados em Organização do Conhecimento, v. 3). p. 575-606.
- GUIMARÃES, J. A. A dimensão teórica do tratamento temático da informação e suas interlocuções com o universo científico da International Society for Knowledge Organization (ISKO). *Revista Ibero-Americana de Ciência da Informação*, Brasília, v. 1, n. 1, p. 77–99, 2008. Disponível em: <https://periodicos.unb.br/index.php/RICI/article/view/940/815>. Acesso em: 01 jul. 2021.
- IFLA. *Functional requirements for subject authority data (FRSAR): a conceptual model*. Den Haag: IFLA, 2010. Disponível em: <http://www.ifla.org/files/classification-and-indexing/functional-requirements-for-subject-authority-data/frsad-final-report.pdf>. Acesso em: 23 jul. 2021.
- IFLA. *Statement of International Cataloguing Principles (ICP)*. Den Haag: IFLA, 2016. Disponível em: <http://www.ifla.org/en/publications/statement-of-international-cataloguing-principles>. Acesso em: 23 jul. 2021.
- KAWAMURA, K. Eric Coates. In: HJORLAND, Birgen; GNOLI, Claudio (ed.). *Encyclopedia of Knowledge Organization*. Toronto: ISKO, 2020. Disponível em: <https://www.isko.org/cyclo/coates>. Acesso em: 29 jun. 2021.
- KAWAMURA, K. Ranganathan and after: Coates' practice and theory. *Advances in Knowledge Organization*, Baden-Baden,

v. 9, p. 337-343, 2004. Disponível em: https://www.ergon-verlag.de/isko_ko/downloads/aikovol09200455.pdf. Acesso em: 29 jun. 2021.

CAPÍTULO 4

REPRESENTAÇÃO DA INFORMAÇÃO DE OBRAS FOTOGRÁFICAS EM AMBIENTE DIGITAL: MUSEU DE ARTE DE SÃO PAULO ASSIS CHATEAUBRIAND (MASP)

Patricia Andreia Lira

Giovana Deliberati Maimone

1 INTRODUÇÃO

A museóloga Waldisa Rússio Camargo Guarnieri (2010) definiu o museu de maneira abrangente, relacionando-o ao seu meio social e ao seu tempo. Guarnieri parte do princípio da relação do museu com a humanidade, como agente histórico, para indicar as modificações que o museu encara para se adaptar à contemporaneidade. Essas modificações estão relacionadas principalmente à comunicação museológica, ou seja, ao acesso público às coleções dos museus.

O desenvolvimento de tecnologias da informação e da comunicação, como explicado por Castells (1999), é promotor de uma nova revolução tecnológica que está em curso e possui a mesma importância que os combustíveis fósseis tiveram em revoluções anteriores. Pode-se dizer que as revoluções dos séculos XVIII e XIX foram tecnológicas e industriais, ao passo que a revolução pela qual passamos é tecnológica e comunicacional. Como afirmam Kranzberg e Pursell (1967

apud Castells, 1999, p. 68) uma das características das revoluções tecnológicas é justamente sua "penetrabilidade, ou seja, sua penetração em todos os domínios da atividade humana, não como fonte exógena de impacto, mas como tecido em que essa atividade é exercida". Dessa forma, o domínio das artes, da cultura, e, portanto, dos museus também é penetrado por essa nova revolução tecnológica.

Considerando os museus como resultado da interação entre o ser humano, sua época e seu meio, e tendo em vista a penetrabilidade da revolução tecnológica em curso, podemos afirmar que os processos de preservação de acervos alteraram-se significativamente com a expansão de instrumentos de trabalho desenvolvidos digitalmente. As atividades de conservação ainda lidam com a manutenção da integridade físico-química de objetos e obras, porém dispõem agora de softwares que auxiliam no controle dos ambientes de guarda, de dispositivos conectados em rede que facilitam a prevenção e o controle de incêndios, e de tecnologias que possibilitam a identificação de intervenções em obras por meio do uso da radiação infravermelha e de raios x. A documentação museológica, isto é, a catalogação, a descrição e a indexação, possui os mesmos objetivos e metodologias anteriores à informatização, porém o uso de sistemas integrados e em rede facilitou a troca de informações entre os profissionais e entre as instituições e o público.

Porém, como afirma Guarnieri (2010) é na comunicação museológica que as tecnologias da informação e da comunicação proporcionaram maiores mudanças. É comum a discussão sobre o momento pós custodial pelo qual os

museus passam, essas instituições encontram-se em uma fase em que à importância dada à conservação física de seus acervos foi adicionada a necessidade de comunicá-los ao público. O acesso e a difusão não se restringem mais às obras em exposição, que costumam ser minoria frente a totalidade dos acervos, eles atingem também as peças em reserva técnica. A ampliação do acesso possibilitada pela digitalização é ainda maior com a disponibilização online desses acervos; para isso foram criados ambientes digitais, localizados nos *websites* dos museus, que permitem a visualização da reprodução digital das obras assim como a consulta às suas informações. Contudo, vale destacar que a informatização da documentação de acervos, assim como a digitalização de peças, obras e documentos, apesar dos avanços recentes, ainda é muito incipiente entre os museus brasileiros, visto que os recursos humanos e orçamentários são escassos e não suficientes para a atualização tecnológica de práticas e procedimentos de preservação e comunicação museológica.

Este capítulo apresenta alguns dos resultados da pesquisa de mestrado desenvolvida no Programa de Pós-Graduação em Ciência da Informação da Escola de Comunicações e Artes da Universidade de São Paulo (ECA-USP). Apresentamos aqui parte dos estudos sobre a representação temática e descritiva da informação relacionada às obras fotográficas disponibilizadas no *website* do Museu de Arte de São Paulo Assis Chateaubriand, o MASP. Estudamos os textos escritos que acompanham as imagens digitais presentes nos *websites*, com o objetivo de verificar se tais textos refletem as necessidades apontadas nas metodologias de

representação da informação imagética desenvolvidas por Manini (2002) e Kossoy (2014).

2 OBRAS FOTOGRÁFICAS NO AMBIENTE DIGITAL DO MASP

Os interessados em conhecer e consultar o acervo do MASP encontram no *website* da instituição um instrumento de pesquisa onde é possível acessar, além de informações sobre a obra consultada, a imagem digitalizada de tal obra. O acesso a esse instrumento de pesquisa ocorre por meio da página “Pesquise no acervo”. Vale ressaltar que o *website* do MASP foi totalmente reformulado em 2017, ocasião na qual ocorreu a integração entre o sistema de gestão do acervo e o novo *website* (MASP, 2017).

A página “Pesquise no acervo” destaca como principais filtros para a recuperação da informação o nome do artista e o título da obra, é possível refinar essa busca pelo ano de produção da obra. O terceiro filtro disponível refere-se à categoria da obra, nesse caso, para realizar sua busca o usuário poderá escolher entre uma das cinquenta e sete opções, incluindo a opção “Todas as categorias”. A lista disponível no *website* do Museu inclui uma gama ampla de categorias possivelmente utilizadas para a classificação do acervo, que vão de termos gerais, como “Fotografia”, “Pintura” e “Desenho”, a específicos, tais como, “Calça”, “Copo” e “Boneco”

O *website* do MASP traz a autoria como o principal ponto de acesso às obras. É segundo a autoria que o acervo é

apresentado ao usuário, é também pela autoria que se dá a organização das obras disponibilizadas ao público: com exceção do título da obra, para qualquer filtro empregado o resultado sempre será apresentado segundo a autoria.

As imagens produzidas a partir da reprodução das obras são evidenciadas na página “Pesquise no acervo”; verificou-se que o retorno de uma determinada busca, utilizando-se qualquer um dos filtros disponíveis, apresenta uma página onde as obras estão organizadas pela autoria, com uma série de imagens relacionadas à busca realizada e expostas na forma de *thumbnails*. Ao clicar na imagem selecionada o usuário é direcionado à outra página, agora com a imagem ampliada seguida por texto descritivo, organizado no formato de ficha, com as informações sobre a obra. Chamaremos essa página de página da obra.

Na página da obra também há o recurso “Trabalhos relacionados”, onde é possível acessar outras obras que possuam relação com o trabalho selecionado, porém tal relacionamento se dá apenas a partir da autoria, ou seja, em “Trabalhos relacionados” pode-se localizar outras obras do mesmo artista cuja ficha foi consultada. Assim temos, novamente, a autoria como fator determinante da classificação do acervo e das escolhas do usuário.

No que compete à fotografia, no momento da pesquisa o *website* do MASP oferecia quatrocentas e treze (413) fotografias para acesso virtual, dessas, apenas cinco trabalhos não possuem imagens em suas páginas da obra. Partindo da verificação das obras designadas como fotografia, pintura, instalação, Assemblage, vídeo, cartaz e desenho, notamos que

as categorias de informação presentes nas fichas que acompanham as imagens na página da obra são sempre as mesmas, a despeito do meio de expressão utilizado pelo artista.

2.1 Representação da informação de obras fotográficas no ambiente digital do MASP

Para compor a amostra do estudo da representação da informação das fotografias presentes na coleção do MASP e acessíveis digitalmente, selecionamos duas obras de seis artistas, totalizando doze fotografias. Os seis artistas foram selecionados tendo em vista a quantidade de fotografias de cada artista acessíveis no site. Os artistas e suas obras selecionadas foram os seguintes: Virginia de Medeiros, *Elizabeth Afonso Pereira, Guerrilheiras*, da série *Alma de Bronze*, 2017 e *Daniela Santos Neves, Guerrilheiras*, da série *Alma de Bronze*, 2017; Ivo Ferreira da Silva, *Elíticas*, sem data e *Círculos de confusão*, sem data; Raul Eitelberg, *Family Portrait*, sem data e *Desesperança*, sem data; German Lorca, *Luiz Hossaka*, 1989 e *Chuva na janela*, 1950; Thomaz Farkas, *Apartamentos ou fachada interior do Edifício São Borja*, ca. 1945 e *Telhas*, ca. 1945; Miguel Rio Branco, *Sem título*, da série *Maciel*, 1979 e *Sem título*, da série *Maciel*, 1979.

Boris Kossoy (2014) considera quatro grupos de informação como os principais para a análise de fotografias com a finalidade de pesquisa, são eles: identidade, assunto, autoria e tecnologia de produção. Data e local de produção são considerados por Kossoy como elementos obrigatórios para representação da imagem fotográfica e estão inseridos nas informações como se verá adiante.

De maneira geral, no que diz respeito à identidade, as fichas apresentadas pelo *website* do MASP trazem as seguintes categorias de informação: “Número de inventário”, “Título” e “Aquisição”. A informação sobre o número de inventário localiza a fotografia em meio ao restante do acervo do Museu, trata-se de uma informação pertinente ao usuário para o caso de solicitações de uso de imagem ou de consulta ao original.

Dado o caráter autoral das fotografias, a informação sobre o título da obra respeita a atribuição definida pelo autor, assim, a amostragem traz a variedade dessas atribuições: Virgínia de Medeiros intitula suas fotos com o nome das retratadas seguida da palavra “Guerrilheiras”, indicando o propósito político das obras; German Lorca também intitula uma de suas fotos selecionadas com o nome do retratado, no caso o fotógrafo Luiz Hossaka; as imagens de Miguel Rio Branco não possuem título; já Raul Eitelberg emprega títulos que simbolizam a atmosfera metafórica e irônica de suas fotografias; a abstração geométrica de Ivo Ferreira da Silva é representada em seus títulos, enquanto que Thomaz Farkas nomeia suas fotos com a identificação direta ao referente. Acompanham os títulos das fotografias de Virgínia de Medeiros e de Miguel Rio Branco os nomes das séries as quais essas fotos pertencem. Em fotografia, série é definida como um conjunto de imagens realizadas a partir de um mesmo objeto, contexto, ideia ou conceito, e podem ser apresentadas em sequência ou descontinuamente. A identificação das séries na ficha da obra indica que aquela fotografia pertence a um determinado grupo originado pela criação do próprio fotógrafo, tal informação, além de ser um elemento

identificador da fotografia, é pertinente para a compreensão de seu contexto de criação.

A aquisição é referenciada com a indicação do tipo, da origem e da data. No caso das fotografias de Virginia de Medeiros, Miguel Rio Branco e do retrato de Luiz Hossaka de autoria de German Lorca o usuário por meio da expressão “doação do artista” obtém a informação sobre o tipo de aquisição (doação) e sua origem (Virginia de Medeiros, Miguel Rio Branco e German Lorca), nesses casos também há a indicação da data da aquisição. As fotografias de Virginia de Medeiros possuem uma informação adicional sobre sua aquisição, elas foram adquiridas no contexto da exposição “Histórias das mulheres, histórias feministas”, dessa forma é possível conhecer parte da trajetória do documento fotográfico, nesse caso em especial a trajetória no Museu. Kossoy (2014) destaca a trajetória do documento como uma informação relevante e digna de registro. As fotografias de Thomas Farkas, Raul Eitelberg, Ivo Ferreira da Silva e *Chuva na janela* de German Lorca foram adquiridas por meio do comodato MASP Foto Cine Clube Bandeirante, assim, ao consultar a ficha sabe-se do tipo de aquisição (comodato) e de sua origem (Foto Cine Clube Bandeirante), porém não há informação sobre a data do comodato.

O próximo grupo de informação tratado por Kossoy (2014) relaciona-se ao assunto da fotografia. Kossoy indica a necessidade de representação do tema da imagem fotográfica por meio de termos descritores de assunto e da descrição concisa da imagem. No entanto, na ficha disponibilizada online

pelo MASP, não há descritores ou mesmo a descrição temática da fotografia.

A autoria possui destaque na metodologia e na teoria desenvolvidas por Kossoy (2014, 2016). Para Kossoy o fotógrafo é determinante no processo fotográfico, pois é baseada em suas escolhas que se dá a construção da imagem. O *website* do MASP, como já citado, prioriza a autoria como filtro para a recuperação da informação, como categoria para a classificação e organização do acervo exibido ao público e como elemento de destaque na ficha da obra, em consequência, a autoria é o principal ponto de acesso ao acervo do Museu através de seu site. As fichas das fotografias selecionadas apresentam o nome do autor e seus dados biográficos, porém os dados biográficos restringem-se à data e ao local de nascimento e morte, quando aplicável, do fotógrafo. Tais dados biográficos são recorrentes não apenas na ficha da obra, mas também na página resultante de um primeiro processo de busca. Aparentemente a função desses dados biográficos está ligada à diferenciação de possíveis homônimas e não à contextualização do período de produção do fotógrafo, apesar de tal informação poder ser deduzida pelo usuário.

A tecnologia empregada na produção da fotografia é outro aspecto apontado na metodologia para a análise de Kossoy (2014). Para a compreensão desse aspecto o autor destaca o registro do equipamento utilizado, do tipo de suporte, do processo empregado, da tonalidade, do formato/dimensões e da montagem. Nas fichas disponíveis no *website* do MASP há duas categorias de informação que

remetem à tecnologia: técnica e dimensões. As dimensões da fotografia original são apresentadas segundo o padrão altura por largura em centímetros, porém não se especifica se tais dimensões referem-se ao suporte ou à mancha da imagem no suporte fotográfico. Já a apresentação da técnica não possui um padrão facilmente identificável: em nossa amostragem, a técnica das sete fotografias adquiridas pelo comodato MASP-FCCB são expostas com a mesma expressão: “Impressão sobre papel prata/gelatina”, indicando tratar-se de fotografias preto e branco, analógicas e em suporte papel. No entanto, a visualização da reprodução da fotografia *Desesperança* de Raul Eitelberg permite-nos constatar que trata-se de uma fotografia colorida, porém não há informação sobre o processo que possibilitou a obtenção da cor (que poderia ser deduzido a partir da investigação dos processos empregados pelo fotógrafo combinada com a data de produção da fotografia e com a visualização do documento original aumentada por microscopia).

Em continuação à representação da informação referente à técnica, as fichas das fotografias de Virgínia de Medeiros apresentam-se como “Fotografia digital, impressão digital sobre papel de algodão”, apesar de informar sobre o processo de obtenção da imagem (“fotografia digital”) e sobre a cópia presente no Museu (“impressão digital sobre papel algodão”), nesse caso também só é possível identificar a cromia/tonalidade da fotografia pela visualização de sua reprodução. Já as fichas das obras de Miguel Rio Branco trazem a seguinte informação: “Fotografia analógica colorida, impressão c-print sobre papel fotográfico”, apesar de

incompleta por não apresentar detalhes como negativo-positivo, diapositivo ou fotografia instantânea, por exemplo, temos a informação sobre o processo fotográfico (“fotografia analógica”) e sobre a cópia do MASP (“impressão c-print sobre papel fotográfico”). Nas fichas das fotografias de Miguel Rio Branco a informação sobre a cromia /tonalidade é duplicada, ela está presente na expressão “Fotografia analógica colorida”, bem como em “impressão c-print sobre papel fotográfico”, onde c-print significa chromogenic print.

A representação da informação sobre a técnica do retrato de Luiz Hossaka realizado por German Lorca, “Fotografia analógica preto e branco, impressão digital sobre papel de algodão”, segue o padrão da representação da técnica das fotografias de Virgínia de Medeiros e Miguel Rio Branco: primeiro indica-se o processo de produção da imagem, mesmo incompleto, (“Fotografia analógica), seguido da informação sobre a cópia do Museu (“impressão digital sobre papel de algodão”).

Por fim, para Kossoy (2014) a identificação do local de produção da fotografia, assim como de sua data de produção, são elementos indispensáveis à análise de fotografias, por esse motivo deveriam estar disponíveis aos usuários que acessam as páginas do “Pesquise no acervo” do *website* do MASP. Porém, em nenhuma ficha há a informação sobre o local de produção das imagens. Em alguns casos, pode-se deduzir o local de produção a partir de outras informações existentes na ficha. Já em relação à data, na maioria dos casos, a amostragem não apresentou a distinção entre a data de

produção das fotografias e a data da cópia preservada pelo Museu, as exceções são as obras de Miguel Rio Branco.

Por sua vez, Manini (2002) propõe o uso da Dimensão Expressiva da imagem fotográfica como um fator para a Análise Documentária de fotografias. A Dimensão Expressiva refere-se aos atributos expressivos específicos da fotografia, escolhidos pelo fotógrafo durante ou após a produção da imagem, tais como o tipo de enquadramento escolhido, os efeitos causados por uso de filtros, a maneira como o contraste é utilizado, a opção por iluminação natural ou artificial etc.

Tal proposta vai ao encontro da necessidade de abranger as especificidades formais da fotografia em sua representação temática, isto é, além do conteúdo da imagem, tal representação poderá elencar aspectos formais indicadores das especificidades da fotografia, esses aspectos estariam representados na descrição e nos termos descritores da imagem fotográfica. Porém, já constatamos que as páginas das obras do *website* do MASP não apresentam a descrição ou mesmo palavras-chave ou descritores temáticos.

Consideramos importante destacar que, apesar de representadas nas fichas da obra, as informações sobre dados biográficos de autor, técnica, dimensões, aquisição e número de tomo não são passíveis de recuperação por meio do sistema de busca disponibilizado no site do Museu.

2.2 A representação temática no ambiente digital do MASP

A ausência da representação temática no *website* do MASP é o ponto que mais nos chamou a atenção na realização

dos estudos. Em sua tese Manini (2002) já indicava essa provável configuração no que ela denominou como “bancos de imagens digitais”. A partir do momento no qual as obras de arte, inclusive a fotografia, são digitalizadas cria-se a possibilidade de disponibilizar as imagens criadas por esse processo de reprodução em instrumentos de pesquisa que encontram na internet um meio de disseminação eficaz e abrangente. Tais instrumentos de pesquisa, agora dotados das imagens das obras de interesse público e favorecidos por diferentes recursos de visualização, acabam por prescindir da representação temática, já que o conteúdo da imagem se encontra acessível e visível.

Consideramos que o objetivo da disponibilização de reproduções digitais de fotografias é ampliar o acesso à imagem; a partir do momento que essa disponibilização se efetua é necessário também oferecer informações para a compreensão da fotografia original, enfim, deve-se esclarecer ao usuário quais são as características do documento original para que ele possa ter instrumentos suficientes para diferenciar a imagem visualizada pelo acesso online da fotografia original, e requerer a consulta ao original caso considere necessário. As características que diferem a imagem digital da fotografia original são referentes à técnica, às dimensões e à data de produção. O museu estudado cumpre esse papel, mesmo que parcialmente: no *website* do MASP encontramos informações referentes à técnica, apesar de não serem apresentados detalhes sobre o processo de produção; às dimensões, com a ressalva da falta de especificação sobre a origem dessas dimensões, se referentes ao suporte ou à

mancha; e sobre a data de produção, a despeito de não existirem informações sobre a origem de tal data, se ela diz respeito ao registro da imagem fotográfica em seu primeiro suporte ou se trata-se da data na qual a cópia fotográfica, o documento presente no Museu, foi produzido. Ao disponibilizar a fonte e a data de aquisição, bem como o número de inventário e tombo, o MASP caracteriza o documento original e favorece a identificação desses documentos em meio aos seus acervos.

A Análise Documentária além de produzir a descrição e a indexação do documento fotográfico, instrumentaliza os sistemas com termos que poderão ser utilizados na recuperação da informação. O MASP não apresenta textos escritos que possam ser considerados como representação temática nas fichas disponíveis em seus *websites*, o Museu tampouco oferece o recurso de localização de obras por palavras-chave temáticas, o principal ponto de acesso é a autoria, o nome do responsável pela produção da obra. Assim, o Museu parte do princípio de que o público que utilizará o *website* para acessar os acervos já conhece o conteúdo das coleções e possui ideia clara sobre a obra que procura. Tal premissa não é de todo falsa, Smit (1987) e Shatford (1986) indicam que a instituição deverá definir o perfil de usuário que irá contemplar ao estabelecer suas práticas de Análise Documentária e de representação da informação, visto que não é possível atender a todas as demandas de todos os tipos de usuários. Portanto, concluímos que o MASP assume como perfil de público pessoas que possuem conhecimento prévio sobre a arte em geral e sobre o acervo da instituição

especificamente. Porém, se é papel do museu na contemporaneidade ampliar e democratizar a produção e a difusão do conhecimento, e considerando que o tratamento do universo informacional presente no museu se destaca na execução desse papel, não seria excludente considerar como público-alvo apenas os já iniciados no campo das artes, tendo em vista, além disso, o uso de um meio que pretende ser inclusivo, como a web? Tal questionamento nos leva a considerar a seguinte frase de Sara Shatford (1986, p. 40): *“We need to understand not only how present users may be better served, but how we might serve those who do not realize that pictures could be useful to them”*.

3 CONSIDERAÇÕES FINAIS

A organização do acervo no *website* do MASP se dá a partir da autoria das obras, já que o nome do autor é o principal ponto de acesso, correspondendo ao papel desempenhado pelos autores nos museus de arte. Crimp (2005) afirma que a legitimação e a autenticação das obras presentes nos museus de arte são dadas, de maneira geral, por suas autorias, afirmação corroborada por Rouillé (2009) ao indicar que a qualidade artística das fotografias é validada por seu autor e não pelo conteúdo ou pela técnica empregada em sua execução. Assim, na maioria dos casos, serão incorporadas aos museus de arte as fotografias cujas autorias sejam reconhecidas e dotadas de certo prestígio. Porém, Crimp (2005) alerta que a primazia da autoria no tratamento dado às fotografias pelos museus de arte poderá frustrar aqueles que a elas recorrem devido ao seu significado histórico ou social,

significado esse que não retira o caráter de obra de arte da fotografia.

Em nossa pesquisa observamos que o contexto histórico e social e a biografia do autor, bem como a descrição temática das obras, podem ser abordados pelos museus ao disponibilizar os acervos em seus *websites*. É recorrente nos sites de alguns dos mais tradicionais museus de arte da Europa e dos Estados Unidos a contextualização do meio social e cultural no qual se deu a produção da obra cuja imagem está em exibição. É usual também a apresentação de relatos sobre a biografia dos artistas. No site do Museo Nacional Centro de Arte Reina Sofia, por exemplo, ao acessarmos a obra *Guernica* (1937) de Pablo Picasso nos deparamos com um texto onde são apresentadas as circunstâncias de realização da obra, sua descrição iconográfica e uma reflexão sobre a relação de Picasso com algumas das figuras representadas. Já o Brooklin Museum traz junto à imagem digital da obra *The Judgement* (O julgamento, 1963) de Bob Thompson um texto breve, mas elucidativo, sobre a carreira do artista e suas influências, o texto é finalizado com a descrição iconográfica da pintura.

Destaca-se a maneira pela qual a representação temática é apresentada nos sites do Musée d'Orsay e da Tate Modern. No Musée d'Orsay, mesmo obras conhecidas e já consolidadas no imaginário popular, como *La Chambre de Van Gogh à Arles* (Quarto em Arles, 1889) de Vincent Van Gogh, é representada tematicamente com o registro de palavras-chave temáticas também utilizadas como *links* para o acesso à outras obras indexadas com os mesmos termos. Já na Tate Modern as palavras-chave estão dispostas como um meio de explorar a

coleção do museu. A representação da obra *Do Women Have To Be Naked To Get Into the Met. Museum?* (As mulheres precisam estar nuas para entrar no Met. Museu? 1989) do grupo Guerrilla Girls no site da Tate é exemplar não apenas pela sua indexação, que apresenta termos relacionados ao conteúdo temático imediato como também aos significados simbólicos, políticos e sociais da obra, mas igualmente pelo texto crítico que discute a importância das artistas para a atualidade da arte e dos museus.

Evidentemente os exemplos expostos acima não esgotam as possibilidades de representação da informação de obras em *websites* de museu de arte, assim como os museus europeus e norte-americanos não podem ser considerados como modelo único de instituição museológica, porém é significativo constatarmos que o contexto de produção, a biografia dos artistas e o conteúdo temático das mais clássicas expressões artísticas, como a pintura e o desenho, são destacados nos sites desses museus. Acreditamos que o mesmo pode ocorrer com a fotografia; o assunto da imagem, seu tema, ou mesmo seu significado, também merecem destaque e enriquecem esses recentes instrumentos de pesquisa, os *websites*.

REFERÊNCIAS

CASTELLS, M. *A sociedade em rede*. São Paulo: Paz e Terra, 1999.

CRIMP, D. *Sobre as ruínas do museu*. São Paulo: Martins Fontes, 2005.

GUARNIERI, W. R. C. Museu, para quê? (A necessidade da arte). In: BRUNO, Cristina (org.). *Waldisa Rússio Camargo Guarnieri, textos e contextos de uma trajetória profissional: a evidência dos contextos museológicos*. São Paulo: Comitê Brasileiro do Conselho Internacional dos Museus, 2010.

KOSSOY, B. *Fotografia & História*. São Paulo: Ateliê Editorial, 2014.

KOSSOY, B. *Realidades e ficções na trama fotográfica*. São Paulo: Ateliê Editorial, 2016.

MANINI, M. P. *Análise Documentária de Fotografias: um referencial de leitura de imagens fotográficas para fins documentários*. 2002. Tese (Doutorado em Ciências da Comunicação) – Escola de Comunicações e Artes, Universidade de São Paulo, São Paulo, 2002.

MASP. *Relatório Anual de Atividades 2017*. São Paulo: Museu de Arte de São Paulo Assis Chateaubriand, 2017. Disponível em: <https://masp.org.br/uploads/about-governance-items/BNYPls0ZiAsWaFNEz4CP2xqg1XB5qVHK.pdf>. Acesso em: 11 out. 2020.

ROUILLÉ, A. *A fotografia: entre documento e arte contemporânea*. São Paulo: Senac, 2009.

SHATFORD, S. Analyzing the Subject of a Picture: A Theoretical Approach. *Cataloging and Classification Quarterly*, Oxfordshire, v. 6, n. 3, p. 39-62, 1986.

SMIT, J. W. A análise da imagem: um primeiro plano. In: SMIT, Johanna W. (coord.). *Análise documentária: a análise da síntese*. Brasília: IBICT, 1987.

CAPÍTULO 5

TEORIA CRÍTICA: UM OLHAR PARA A ORGANIZAÇÃO DO CONHECIMENTO NA WEB

Luciane Paula Vital

1 INTRODUÇÃO

A área de Organização do Conhecimento (OC) busca discutir teorias e metodologias que viabilizem a análise, compreensão, representação e organização de áreas específicas do conhecimento humano. Pensando do ponto de vista do desenvolvimento do conhecimento, de diferentes naturezas, evidencia-se a complexidade da representação e organização abarcar em um sistema único essas visões, que em alguns casos, antagonizam. A área de OC tem investido fortemente em estudos que apresentam, discutem e aprofundam a diversidade que os Sistemas de Organização do Conhecimento (SOC) precisam considerar em seu desenvolvimento. Esse fato pode ter como força propulsora a rede de computadores, que, com as tecnologias da informação e comunicação, tem conectado pessoas ao redor do mundo em torno de temáticas de interesse comum, mas não necessariamente com pontos de vista convergentes.

A facilidade proporcionada pelo Web de disponibilização e acesso às informações de diversas naturezas, também apresenta o desafio de uma representação e organização que deem conta de abarcar as diferenças, com espaço para concepções de mundo diversas. A União

Internacional de Telecomunicações, agência da Organização das Nações Unidas (ONU) especializada em tecnologias de informação e comunicação, estimava que em 2019 mais de 3 bilhões de pessoas ainda não teriam acesso à internet, a maioria são mulheres (52%), contra 42% dos homens, além de serem originários de países menos desenvolvidos economicamente (ONU, 2019). Ou seja, essas pessoas não estão participando do desenvolvimento da Web e, conseqüente, não estão sendo representadas nesse espaço. Nesse sentido, é fundamental a participação de grupos heterogêneos no desenvolvimento de SOC, criando espaço para que as diferenças sejam postas e compartilhadas de forma a facilitar o acesso e recuperação de informações por todos, de forma democrática.

A Web é um dos campos de poder das relações humanas, onde prevalecem grupos privilegiados historicamente, homens, brancos, europeus, heterossexuais, por exemplo, com suas visões de mundo. Para que as instituições de informação não perpetuem visões historicamente privilegiadas, reproduzindo uma representação e organização pautada nessa lógica dominante, é importante um olhar crítico sobre o processo de desenvolvimento dessas representações. Apresentamos e refletimos sobre as possibilidades que a teoria crítica tem no desenvolvimento de estudos que contemplem pontos de vista de grupos sociais heterogêneos na OC e, conseqüentemente, na construção de SOC, especialmente na Web, pela facilidade de disseminação e impacto do que é produzido nesse ambiente.

2 TEORIA CRÍTICA

Não é pretensão desse texto abordar exaustivamente as discussões que permeiam a teoria crítica, a intenção é caracterizá-la brevemente buscando conectá-la e problematizá-la como possibilidade de olhar a OC, especialmente no contexto da Web. Já que a teoria crítica intenta a conexão com o mundo real e atual, contextualizando-se histórica e geograficamente, quando desenvolvemos SOC no ambiente digital, dada a heterogeneidade posta, ela pode nos fornecer elementos que qualifiquem esse processo de forma democrática e como instrumento emancipatório.

O aparecimento do termo “teoria crítica” para demarcar estudos com viés crítico social tem como marco o texto de Max Horkheimer, de 1937, intitulado “Teoria tradicional e teoria crítica”, publicado pela revista do *Institut für Sozialforschung* (Instituto de Pesquisas Sociais), vinculado à Universidade de Frankfurt, sete anos após o autor assumir a direção do instituto. Horkheimer marca o que alguns autores definem como a primeira geração de autores da teoria crítica, que desenvolviam estudos vinculados ao que foi denominada como 'Escola de Frankfurt'. A reconstituição de fatos históricos que marcaram o início do pensamento crítico social nos interessa menos que os fatores que levaram esses pensadores a buscarem uma postura crítica ao pensarem seus objetos de pesquisa e, conseqüentemente, a sociedade na qual viviam. Nem mesmo o texto de Horkheimer (1937 *apud* Fleck, 2017) define o conceito de teoria crítica, talvez propositadamente para não estabelecer limites a algo que está(va) se delineando. Fleck (2017, p. 111) afirma que,

[...] uma das coisas que certamente se pode falar desta tradição teórica é que suas reflexões não tratam apenas do conteúdo propriamente dito (a sociedade em seus múltiplos aspectos), mas também da forma como este conteúdo é apreendido. A teoria crítica sempre foi, também, uma espécie de metateoria. [...] Assim, a relação entre sujeito e objeto é diferente na teoria tradicional e na teoria crítica. Enquanto um abismo separa um do outro na tradicional, sujeito e objeto se codeterminam na crítica, de forma que o sujeito se torna sujeito por meio de sua confrontação com o objeto (e o mesmo pode ser dito da relação entre forma e conteúdo).

A teoria crítica apresenta base marxista, como crítica ao positivismo e a ideologia capitalista de se render ao mundo da forma como as relações estão postas, visa uma orientação à emancipação. Para Nobre (2004, p. 11),

[...] a teoria crítica não pode se confirmar senão na prática transformadora das relações sociais vigentes. As ações a serem empreendidas para a superação dos obstáculos à emancipação constituem-se em um movimento da própria teoria. Nesse sentido, o curso histórico dos acontecimentos — como resultado das ações empreendidas contra a estrutura de dominação vigente — dá a medida para a confirmação ou refutação dos prognósticos da teoria.

Isso significa, na perspectiva da teoria crítica, uma mudança de postura, de questionar a estrutura vigente e seus olhares sobre ela e os objetos do conhecimento que a constituem. Inclui questionar e criticar sua própria postura diante do seu objeto de análise, qual a concepção de mundo e o ponto de vista do qual analiso o mundo que me cerca. O

teórico crítico canadense, Andrew Feenberg, apresenta o que denomina 'Teoria crítica da tecnologia', que insere a teoria crítica no contexto das tecnologias da informação e comunicação,

O projeto geral de Feenberg será o de buscar destacar um conceito de essência da técnica enquanto fenômeno tanto imanentemente técnico como social, histórico e político. Buscará, portanto, os estudos empíricos que revelam e debatem como os interesses de certos atores alcançam e empreendem o controle dos processos de escolha, concepção e design das materialidades tecnológicas; e como a concretização dessas últimas agenciam novos valores e decisões na sociedade. (Genaro, 2017, p. 297)

A técnica nunca se constitui somente em técnica, é constituída do contexto político, econômico, cultural e social que a circunda. Compreendemos que um olhar crítico, atento aos processos de OC requer rever, discutir, analisar os posicionamentos adotados no SOC de forma perene, partindo da premissa que são valores e decisões, como pontua Feenberg (2013), que afetam o desenvolvimento da vida em sociedade. Se a queremos democrática, é preciso iniciar pelo questionamento das atividades que como profissionais da informação desenvolvemos.

3 ORGANIZAÇÃO DO CONHECIMENTO NA WEB

Na área da Organização do Conhecimento é possível identificar, segundo autores como Melo (2013) e Francelin e Kobashi (2011), diferentes abordagens nos estudos sobre o

conceito, sendo duas principais: a pragmática, sendo um dos representantes Birger Hjørland, e a positivista, representada por Ingetraut Dahlberg.

Melo (2013) discute as características dessas duas abordagens que influenciam a epistemologia da OC. O positivismo, por buscar uma verdade absoluta e livre das interferências do meio, "[...] admite a possibilidade de uma organização ideal do conhecimento por meio da representação objetiva do significado (essência) de um termo." (Melo, 2013, p. 37). Essa objetividade lógica é possível verificar na definição de conceito apresentada por Dahlberg (1978, p. 102), "É possível definir, então, o conceito como a compilação de enunciados verdadeiros sobre determinado objeto, fixada por um símbolo linguístico.", essa verdade independe do contexto no qual o conceito está sendo formulado, admitindo a universalidade. Ainda, segundo Woods (1975 *apud* Welty, 1996), modelos de representação do conhecimento seriam aplicados na representação inequívoca de alguma interpretação de uma sentença (lógica), com um método para 'tradução' da linguagem natural para a de representação. Percebe-se nessas definições uma visão estruturada do mundo, onde objetos e conceitos podem ser delimitados, definidos, representados e organizados de forma objetiva e independente do contexto.

Já o pragmatismo, ainda de acordo com Melo (2013, p. 42),

[...] promove um reencontro do sujeito com o objeto de estudo da Ciência da Informação. Compreende que somente uma interpretação

profunda do homem em suas relações sociais possibilita analisar e organizar a construção coletiva do conhecimento.

A partir desse entendimento, Hjørland (2009, p. 1530, tradução nossa) diz que os "conceitos tem sido entendidos como significados socialmente negociados que podem ser identificados por estudos do discurso em vez de estudar usuários individuais ou princípios *a priori*." Essa afirmação de Hjørland nos remete à relação sujeito-objeto posta na teoria crítica, uma construção que se faz na interação e transforma ambos. O contexto, sujeito e objeto e suas diferentes implicações podem ser considerados, nessa perspectiva, como elementos centrais na construção de SOC.

Como afirmamos, o campo da Web é o campo de poderes, as representações se dão de acordo com fatores externos, muitas vezes, alheios à vontade do indivíduo. Harari (2016) chama de dataísmo a 'religião dos dados', que para ele representa o protagonismo que os dados têm recebido na atualidade. O autor aborda a capacidade humana de enriquecer dados, informações para que se transformem em conhecimento, no contexto do dataísmo, considerando essa uma tarefa árdua, com a produção de grandes volumes de dados. Esse fato torna essa capacidade humana dependente de sistemas computacionais, o que criaria uma dependência em relação às máquinas. A possibilidade de processamento e análise de diferentes tipos de dados pelas mesmas ferramentas, segundo Harari (2016, p. 371), é uma ideia

extremamente atraente. Ela oferece a todos os cientistas uma linguagem comum, constrói pontes sobre brechas acadêmicas e exporta

facilmente insights através das fronteiras disciplinares.

A transformação do mundo nesse grande processador de dados, seguindo o questionamento de Harari (2016), nos faria chegar a um consenso sobre qual o resultado mais acertado desse processo? Existiria uma representação ideal e universal proposta pelos algoritmos, conforme já preconizou Paul Otlet? Essas questões são importantes para nos situarmos no tempo-espaço da cultura dos dados em que nos encontramos, quais serão as consequências dessas práticas ou qual será o papel do ser humano nesse contexto. A área de OC no Brasil já vem questionando essa possibilidade por meio de pesquisas que abordam o contexto cultural, social e político dessa organização (Vital; Hernandez; Santos; 2019). O contexto brasileiro é fortemente influenciado por um movimento internacional da OC, como vimos anteriormente, pragmatismo, que questiona e insere as questões da organização a partir de um viés que contempla as diferenças, que busca 'resgatar' o olhar do sujeito marginalizado na sua relação com o objeto em áreas específicas do conhecimento.

O contexto da Web desafia a área de OC, se constitui em uma arena que se diz única, em que todas as conexões ocorrem, em que o fluxo de dados e informações é livre, em que as diferentes 'tribos' se encontram. Mas, esse encontro é democrático de fato? Podemos perceber a diversidade sendo representada nesse espaço tão central em nossas vidas cotidianas? Garcia Gutierrez, pesquisador da área de OC, pontua, em texto de 2011, que a rede digital coloca o Ocidente em evidência, ditando as regras desse ambiente, de forma

metonímica, dicotômica e neocolonial. O autor propõe um reposicionamento da teoria e da prática da OC para contemplar o mundo heterogêneo em que estamos inseridos por meio de abordagens críticas. O reposicionamento proposto é teorizado no que Garcia Gutierrez (2011, p. 13) chama de operadores desclassificantes, chamados complexo e transcultural.

Under the priority of ideological and logical pluralism, but also facilitating part of that established in the principles of interaction and transformation that orientate the promotion of social change, I propose complex operator Λ . [...] Operator Λ would guarantee all the ideological interpretations and equal opportunities of those concepts. The complex operator is not designed to intervene in or control visions and meanings as regards an issue.¹

A ideia de Garcia Gurierrez é a de que o operador complexo Λ detecte todas as visões marginalizadas, não vistas e/ou consideradas, para que possam ser trazidas para o processo de organização e representação do conhecimento por meio do operador transcultural V .

With regard to transcultural operator V , it is the synthetic product of a permanently open, democratic dialogue between representations of diverse positions (political, cultural, discursive,

¹ Sob a prioridade do pluralismo ideológico e lógico, mas também facilitando parte do estabelecido nos princípios de interação e transformação que orientam a promoção da mudança social, proponho o operador complexo Λ . [...] O Operador Λ garantiria todas as interpretações ideológicas e igualdade de oportunidades desses conceitos. O operador complexo não foi concebido para intervir ou controlar visões e significados relativos a um problema. (Garcia Gutierrez, 2011, p. 13, tradução nossa).

etc.) that negotiate the homologation and integration of certain issues that affect them from argument premises (and not mere arguments) or topoi. (Garcia Gutierrez, 2011, p. 15).²

Falta a Enquanto o operador complexo apresenta as múltiplas facetas de uma área do conhecimento, e suas relações conceituais, por exemplo, o operador transcultural busca o consenso, a regulação sobre qual a representação adequada no contexto, porém, mostrando as outras possibilidades de compreensão sobre o mesmo objeto. Configuram-se em operadores que buscam uma discussão e regulação democrática do processo de OC. Nas palavras de Garcia Gutierrez (2011, p. 15),

Both operators are profoundly democratic, since when / is supported by the specification of all the positions and worldviews, without exclusion, as itineraries of representation and localization of the registers, that is, guaranteeing the representativeness on an equal footing of all the initial positions with respect to an issue, V is essentially regulative and executive, that is, it balances the possible unjust treatment of some or other position in the network, even respecting the presence of such registers accommodating the prior principle of emancipation, in terms of a transculturally accepted categorical scale in such a way that abuses in the network do not go

² No que diz respeito ao operador transcultural V, ele é o produto sintético de um diálogo democrático permanentemente aberto entre representações de posições diversas (políticas, culturais, discursivas etc.) que negociam a homologação e integração de determinadas questões que as afetam a partir de premissas argumentativas (e não meros argumentos) ou topoi. (Garcia Gutierrez, 2011, p. 15, tradução nossa).

*unpunished if the transcultural community can avoid it with alerts, warnings and recriminations.*³

É interessante observar que esse reposicionamento proposto por Garcia Gutierrez (2011) considera que as abordagens são múltiplas, algumas tem espaços privilegiados e que precisam ser questionados, mas, o processo de OC sempre fará uma escolha deliberada por um desses aspectos no momento da organização e representação. A proposta do autor é a de que outras abordagens, entendimentos e/ou pontos de vista encontrem espaço também, de alguma forma sejam apresentados na representação, deixando ao usuário a possibilidade de escolha. Operacionalmente não é um processo simples, considerando a perspectiva de recuperação da informação, mas a proposição fomenta uma discussão fundamental na área, que ganha maior relevância no ambiente Web, que é dar espaço à diversidade que esse ambiente apresenta.

Nesse sentido, a teoria crítica também abre um espaço interessante e importante de discussão para a OC. Martínez-Ávila, Semidão e Ferreira (2016) já analisaram a teoria crítica

³ Ambos os operadores são profundamente democráticos, desde quando se apoia na especificação de todas as posições e visões de mundo, sem exclusão, como itinerários de representação e localização dos registros, ou seja, garantindo a representatividade em pé de igualdade de todas as posições iniciais no que diz respeito a uma questão, V é essencialmente regulador e executivo, ou seja, equilibra o possível tratamento injusto de uma ou outra posição na rede, respeitando a presença de tais registros que acomodem o princípio prévio da emancipação, em termos de uma escala categórica transculturalmente aceita de tal forma que os abusos na rede não fiquem impunes se a comunidade transcultural puder evitá-los com alertas, advertências e recriminações. (Garcia Gutierrez, 2011, p. 15, tradução nossa).

como uma resposta às questões éticas verificadas na OC, o que indicia uma busca da área por bases teóricas e metodológicas alinhadas com perspectivas que deem vez e voz à diversidade e grupos marginalizados. Os autores analisaram trabalhos que se caracterizam na vertente da teoria crítica e a influência dessa escolha nos procedimentos metodológicos adotados. Essa questão posta pelos autores é central, se quisermos uma OC plural, precisamos pensar nas metodologias que guiam as pesquisas desenvolvidas na área. Então se coloca a seguinte questão, de que forma desenvolver a organização e representação do conhecimento realizada por sujeitos historicamente situados, imbricados na relação sujeito-objeto, de maneira que contemple as múltiplas visões de mundo, as diferentes facetas de análise possíveis? Reiterando que um dos objetivos para a construção dos SOC é que sejam guias para o entendimento e a recuperação em áreas do conhecimento humano, constituindo-se em uma grande e importante tarefa, que abrange o coletivo da sociedade. Martínez-Ávila, Semidão e Ferreira (2016) dizem que a metodologia na teoria crítica combina ética, a busca por justiça social e valores/elementos pragmáticos do ponto de vista de algum grupo social que não é considerado nos sistemas universais de classificação, por exemplo. Dessa forma, para pensarmos o desenvolvimento de SOC é premente abrir espaço para a participação dos sujeitos envolvidos nessa representação de forma que sejam acolhidas suas demandas. Populações marginalizadas, países que não estão no centro do desenvolvimento econômico mundial, regiões discriminadas dentro de um país ou continente, e, pensando localmente, comunidades periféricas, precisam ser considerados e representados nos processos de OC. Essa ação

pode facilitar o acesso ao conhecimento, que é base para a emancipação dessas pessoas, o que parece ser a parcela da OC no desenvolvimento de uma Web mais democrática para um mundo mais ético.

4 CONSIDERAÇÕES FINAIS

Partindo das reflexões trazidas pela teoria crítica podemos pensar a área de OC e seu desenvolvimento no contexto da Web. Sendo os ambientes digitais tão centrais na atualidade e espaços de poder muito importantes, que marginalizam uma parcela significativa da população, como profissionais da informação e desenvolvedores de SOC, precisamos questionar a forma com que nossas atividades estão sendo desenvolvidas. Estamos dando espaço para a expressão da diversidade? Pensamos em quais grupos estão sendo privilegiados e quais estão sendo marginalizados no processo de OC que desenvolvemos? Nossas pesquisas têm buscado formas de abrir espaço para a participação de grupos heterogêneos? São algumas das questões possíveis de se fazer nesse processo, como já salientamos, potencializados no ambiente Web.

A teoria crítica se apresenta como uma possibilidade teórica e metodológica para subsidiar estudos na área de OC, como já vêm sendo citada e utilizada em pesquisas, ainda de forma limitada. Uma das contribuições mais relevantes da teoria crítica, em nosso ponto de vista, é a sinalização de que precisamos estar alertas em cada etapa do desenvolvimento de representações em busca de constituí-las em espaços de

encontro e negociação de visões heterogêneas, partindo dos usuários dos SOC.

REFERÊNCIAS

DAHLBERG, I. Teoria do conceito. *Ciência da Informação*, Brasília, v. 7, n. 2, p. 101-07, 1978.

FEENBERG, A. A tecnologia pode incorporar valores? A resposta de Marcuse para a questão da época. In: NEDER, Ricardo T. (org.). *A teoria crítica de Andrew Feenberg: racionalização democrática, poder e tecnologia*. 2. ed. Brasília: Observatório do Movimento pela Tecnologia Social na América Latina / CDS / UnB / Capes, 2013. Cap. 9. p. 290-336. Disponível em: <https://www.sfu.ca/~andrewf/coletanea.pdf> Acesso em: 29 maio 2024.

FLECK, A. Afinal de contas, o que é teoria crítica? [After all, what is critical theory?]. *Princípios: Revista de Filosofia*, Natal, v. 24, n. 44, p. 97-127, 21 ago. 2017.

FRANCELIN, M. M.; KOBASHI, N. Y. Concepções sobre o conceito na organização da informação e do conhecimento. *Ciência da Informação*, Brasília, v. 40, n. 2, p. 207-228, maio/ago., 2011.

GARCIA GUTIERREZ, A. Desclassificação na organização do conhecimento: ensaio pósepistemológico. *Transinformação*, Brasília, v. 23, n. 1, p. 5-14, 2011.

GENARO, E. de. O debate da Teoria Crítica sobre a tecnologia. *Ciências Sociais Unisinos*, São Leopoldo, v. 53, n. 2, p. 292-299, maio/ago. 2017.

HARARI, Y. N. *Homo Deus: uma breve história do amanhã*. São Paulo: Companhia das Letras, 2016.

HJORLAND, B. Concept theory. *Journal of the American Society for Information Science and Technology*, New York, v. 60, n. 8, p. 1519-1536, 2009.

MARTÍNEZ-ÁVILA, D.; SEMIDÃO, R.; FERREIRA, M. Methodological Aspects of Critical Theories in Knowledge Organization. *Knowledge Organization*, Baden-Baden, v. 46, n. 2, p. 118-125, 2016.

MELO, M. A. F. *Contribuições das abordagens positivista e pragmática do estudo do conceito para o modelo conceitual FRASAD*. 2013. Dissertação (Mestrado em Ciência da Informação) – Universidade de Brasília, Brasília, 2013.

NOBRE, M. *Teoria crítica*. São Paulo: Jorge Zahar Ed., 2004. (Coleção Filosofia passo-a-passo; 47).

ONU. *Estudo da ONU revela que mundo tem abismo digital de gênero*. 2019. Disponível em: <https://news.un.org/pt/story/2019/11/1693711>. Acesso em: 10 jul. 2021.

VITAL, L. P.; HERNANDEZ, B. F.; SANTOS, A. dos. Análise de conteúdo da produção científica da Organização do Conhecimento em sua dimensão política e social no Brasil. *Informação & Informação*, Londrina, v. 24, n. 3, p. 1–27, 2019. DOI: 10.5433/1981-8920.2019v24n3p1. Disponível em: <https://ojs.uel.br/revistas/uel/index.php/informacao/article/view/38709>. Acesso em: 20 maio 2024.

WELTY, C. A. *An integrated representation for software development and discovery*. 1995. Tese (Doutorado em Ciência da Computação) – Instituto Politécnico Rensselaer, Nova York, Estados Unidos, 1995.

CAPÍTULO 6

HIBRIDIZAÇÃO DAS FOLKSONOMIAS E INSTRUMENTOS DE CONTROLE TERMINOLÓGICO PARA A REPRESENTAÇÃO DO CONHECIMENTO NA WEB: TENDÊNCIAS E DESAFIOS

Raimunda Fernanda dos Santos

1 INTRODUÇÃO

Com o advento de aplicações da *web* que facilitam a participação ativa dos usuários, as atividades de representação de conteúdos nesse cenário têm sido realizadas por sujeitos e máquinas por meio das Folksonomias. Através dela os usuários (humanos ou robôs) assumem o protagonismo nas práticas de representação e recuperação dos objetos informacionais em ambientes colaborativos, possibilitando que, através da linguagem natural, sejam selecionadas e atribuídas as palavras-chave consideradas mais pertinentes para descrever itens informacionais, sejam eles áudios, textos, imagens, vídeos, etc.

Nessa perspectiva, as Folksonomias são consideradas como o resultado da atribuição livre de etiquetas (também denominadas de *tags*, palavras-chave) realizada por humanos ou robôs em sistemas colaborativos para fins de representação e recuperação de objetos informacionais.

As Folksonomias vêm se configurando em um campo de estudo ainda a ser explorado na Ciência da Informação, principalmente no que concerne às diversas ferramentas, modelos, serviços, sistemas e aplicações em que elas podem estar inseridas. Elas apresentam potencialidades para a construção, atualização e hibridização de Sistemas de Organização do Conhecimento, ao mesmo tempo em que também podem ser enriquecidas semanticamente por instrumentos de controle terminológico em sistemas colaborativos.

Os usuários atribuem sentido às coisas pela maneira com as quais utilizam ou as integram e associam em suas práticas cotidianas; bem como pela forma que as representam por meio das palavras empregadas que, posteriormente, são classificadas e conceituadas. Nessa perspectiva, concorda-se com Limpens, Gandon e Buffa (2010), quando os mesmos ressaltam que as Folksonomias consistem em um sistema colaborativo de organização do conhecimento que contempla uma estrutura triádica: etiquetas, pessoas e recursos.

Através das Folksonomias é possível que usuários (humanos ou robôs) atribuam etiquetas ou *tags* aos recursos informacionais que julgam melhor representar as informações contidas nesses itens. Essa prática possibilita a aproximação entre os usuários e os itens informacionais, além de ampliar as chances de encontrabilidade e as possibilidades de busca e navegação, por meio de *tags* associadas às características e às especificidades desses recursos.

A sinergia das Folksonomias e dos instrumentos de controle terminológico pode trazer benefícios para os sistemas

colaborativos de representação e recuperação da informação no contexto da web semântica, extensão da web atual cujos conteúdos recebem significados bem definidos- possibilitando que as pessoas e máquinas possam trabalhar em cooperação com vistas a recuperação semântica dos objetos informacionais.

Essa cooperação resulta em sistemas híbridos de organização do conhecimento, os quais agregam as potencialidades das Folksonomias e de outro(s) Sistemas de Organização do Conhecimento (ontologia, taxonomia, tesouro, por exemplo), podendo ser aplicados em domínios específicos, objetivando a produção significativa de metadados semânticos para fins de representação e recuperação de conteúdos na web.

Diante da capacidade e versatilidade das Folksonomias para integrar outros Sistemas de Organização do Conhecimento, bem como da possibilidade de ser enriquecida semanticamente por instrumentos como ontologias, taxonomias e tesouros, surge o seguinte questionamento: Como as Folksonomias e os instrumentos de controle terminológico podem contribuir simultaneamente para a representação do conhecimento em sistemas colaborativos?

Visando responder esse questionamento, esta pesquisa tem como objetivo geral analisar questões concernentes à hibridização das Folksonomias e instrumentos de controle terminológico para a representação do conhecimento em cenários colaborativos.

Para o desenvolvimento deste estudo foram realizadas as pesquisas bibliográfica, exploratória e descritiva com abordagem qualitativa a partir da coleta e análise dos aspectos enfatizados nas produções científicas nacionais e internacionais sobre garantias que subsidiam a construção de SOC. Além disso foram recuperadas produções científicas sobre a integração das Folksonomias e instrumentos de controle terminológico (como ontologias, taxonomias e tesouros). Para tanto, foram consultadas as seguintes fontes de informação: Portal de Periódicos CAPES; Base de Dados em Ciência da Informação (BRAPCI) e Biblioteca Digital Brasileira de Teses e Dissertações (BDTD). Os trabalhos recuperados nessas fontes foram citados como referencial teórico desta pesquisa com o objetivo de apresentar ideias e reflexões que possam colaborar com as contribuições deste estudo.

A importância desta investigação decorre, em linhas gerais, da necessidade de estudar as novas possibilidades de representação e recuperação de conteúdos que conciliam as bases folksonômicas e de controle terminológico para apoiar a produção significativa de metadados semânticos na web.

As *tags* podem ser analisadas sob o viés dos princípios éticos, de uso, culturais e semânticos, os quais também se encontram ancorados nas garantias consideradas no processo de construção e atualização dos instrumentos de controle terminológico.

Seguindo essa linha de raciocínio, a seguir são apresentadas questões concernentes aos princípios que servem como base para a construção e atualização de instrumentos de controle terminológico, os quais carecem de

ser estudados sob o viés das práticas folksonômicas enquanto representação social do conhecimento dos indivíduos.

2 PRINCÍPIOS PARA A CONSTRUÇÃO DE SISTEMAS DE ORGANIZAÇÃO DO CONHECIMENTO

Os aspectos culturais, éticos, semânticos e de uso também assumem relevância nas discussões acerca das práticas de indexação e de construção de Sistemas de Organização do Conhecimento (SOC), esses últimos expressam valores que podem ser investigados à luz das garantias sobre as quais são desenvolvidos.

Sobre esse prisma, Beghtol (1986) enfatiza que o conceito de garantia está associado às noções de autoridade, justificação e validação. Essa tríade, por sua vez, deve ser contemplada nas diretrizes que apontam quais conceitos devem ser inseridos em instrumentos de controle terminológico, bem como as características de sua categorização e ordenação.

Para Lancaster (2004) esses instrumentos consistem em listas de termos ou descritores controlados, construídas para fins de tratamento e recuperação da informação objetivando controlar sinônimos, diferenciar homógrafos e agregar termos cujos significados apresentem alguma relação mais estreita entre si. Diante disso, observa-se que esses instrumentos objetivam controlar as dispersões lexicais existentes nas linguagens naturais; orientar o indexador no que concerne à escolha dos termos para representar o assunto

dos materiais informacionais e indicar os melhores termos para a indexação.

Com a ampliação da quantidade de sistemas de recuperação da informação e o incremento de tecnologias aplicadas às atividades de representação, organização, recuperação, acesso e uso da informação, surgiram instrumentos complementares aos sistemas de classificação como as listas de cabeçalho de assunto, taxonomias, tesouros, ontologias e Folksonomias, por exemplo. Esses instrumentos têm sido denominados genericamente por *Knowledge Organization Systems* (KOS), Sistemas de Organização do Conhecimento (SOC), em português. De acordo com Soler Monreal e Gil Leiva (2010), eles se diferenciam de acordo com o grau de complexidade e formalização semântica dos termos, possuindo características funcionais e estruturais peculiares.

As garantias subsidiam a compreensão, construção e atualização de Sistemas de Organização do Conhecimento e contemplam um conjunto de preceitos para a validação e inclusão de consensos semânticos nesses sistemas. As pesquisas sobre essa temática despontaram a partir dos estudos sobre garantia literária. De acordo com Zamboni (2018) o termo “garantia” foi inicialmente utilizado por E. Hulme com vistas a descrever a “garantia literária” por meio da noção de que os descritores constituintes de um sistema de classificação devem ser eleitos a partir da literatura a ser classificada.

Para Barité (2007) o princípio da garantia literária pode ser aplicado em diferentes perspectivas como ferramenta de validação e de identificação de conceitos centrais e periféricos

de um domínio; de qualificação para a recuperação do conteúdo de informações especializadas e de avaliação de linguagens documentárias e estruturas de conceitos.

Tendo em vista as limitações da garantia literária e a importância de integrar outros aspectos metodológicos relacionados às perspectivas diferentes daquela apresentada por Hulme, outros modelos de garantias foram elaborados, como por exemplo: garantia cultural, garantia organizacional, garantia ética, garantia científica, garantia semântica, etc. Essas formas de garantia geralmente são usadas em combinação e foram elaboradas com vistas a melhorar a qualidade dos Sistemas de Organização do Conhecimento. Para tanto, não podem ser interpretadas como subordinações ou tipologias dependentes da garantia literária, elas podem ser consideradas como derivações da mesma (Barité, 2011).

A norma Z39.19-2005 da NISO (*National Information Standards Organization*, Organização Nacional de Padrões da Informação) inclui dois outros tipos de garantia, além da garantia literária, são elas: a garantia organizacional e a garantia de usuário. A inclusão da garantia organizacional é justificada pela importância da análise e seleção dos termos considerando o contexto da organização. De outro modo disposto, a garantia de usuário está pautada na importância da seleção dos termos por sua repetição nas buscas realizadas pelos usuários. Essa norma também inclui a garantia cultural, todavia essa tipologia de garantia é considerada somente em vocabulários multilíngues.

A garantia semântica, por sua vez, tem como objetivo orientar o desenvolvedor a verificar, justificar e validar os

significados das unidades terminológicas em classes e conceitos de um SOC para uma determinada finalidade ou audiência (Guedes, 2017).

Com isso, verifica-se a preocupação em ampliar, por meio de perspectivas interdisciplinares, a noção de garantia para diversos elementos teórico-metodológicos que contribuem para a organização e representação do conhecimento de maneira colaborativa. Afinal, a noção de garantia pode ser associada a uma “[...] lente através da qual podemos avaliar na medida em que uma classificação tem sentido e utilidade para os propósitos para os quais foi criada” (Kwasnik, 2010, p. 107, tradução nossa).

Os aspectos relacionados à cultura também assumem relevância no contexto das garantias, principalmente quando se trata de discussões sobre valores compartilhados, história e memória coletiva, por exemplo. Segundo Beghtol (1986) o termo “garantia cultural” foi cunhado por J. M. Lee ao discutir a visão de que “a base semântica ou categorial de uma dada classificação é o produto da cultura que a produziu”.

Para Beghtol (2002, p. 45, tradução nossa) a garantia cultural “[...] pressupõe que todo sistema de classificação é baseado nas premissas e preocupações de uma determinada cultura, seja a cultura de um país, ou uma unidade social maior ou menor (exemplo: grupo étnico, disciplina acadêmica, domínio, partido político, religião e/ou língua)”, considerando que indivíduos que fazem parte de diferentes culturas necessitam de diferentes tipos de acesso à informação.

Esse aspecto de complexidade e, ao mesmo tempo, de diversidade está associado às variadas significações, interpretações, representações existentes na própria cultura, pois para Hall (2016, p. 20) “em toda cultura há sempre uma grande variedade de significados a respeito de qualquer tema e mais de uma maneira de representá-lo ou interpretá-lo”, tendo em vista que a cultura se relaciona a conceitos, ideias, sentimentos, emoções e senso de pertencimento.

De certa forma, “a garantia cultural introduz o fator ético na Organização do Conhecimento” (Barité, 2011, p. 6, tradução nossa). Esse fator se torna elemento importante na atualidade, dada a intensificação do contato entre as culturas, o qual é facilitado pelas Tecnologias de Informação e Comunicação. Com isso, para fins de construção e atualização de Sistemas de Organização do Conhecimento, se faz necessário refletir acerca dos desafios éticos que se tornam, gradativamente, socialmente mais complexos.

Seguindo essa linha de raciocínio, visando resolver questões relacionadas a esses desafios éticos, Beghtol (2002) propõe o emprego do conceito de hospitalidade cultural, estabelecendo conexões entre a garantia cultural e a garantia ética. Esse termo diz respeito à capacidade de um SOC incluir conceitos desconsiderados, de maneira intencional ou não, acomodando várias garantias e culturas, no intuito de refletir, de maneira apropriada, as crenças de quaisquer indivíduos, grupos ou comunidades (Beghtol, 2005, p. 905).

Os indivíduos apresentam significados a objetos, pessoas e eventos por meio de paradigmas de interpretação direcionados a eles, bem como pela maneira como eles as

utilizam ou as integram e associam em suas práticas cotidianas. Nas concepções de Francelin e Pinho (2011) essas práticas são materializadas por meio da terminologia proveniente dessa comunidade discursiva ou área do saber. Por isso, faz-se necessário realizar estudos terminológicos por meio da análise de domínio ou de comunidades discursivas.

Somado a isso, considera-se também importante a realização de investigações em ambientes colaborativos cujas práticas de representação de conteúdos são realizadas por diferentes usuários através das Folksonomias.

As práticas de indexação e classificação devem contemplar as questões culturais, sobretudo pelo discurso em que uma comunidade ou grupo busca solidificar a sua identidade.

Nesse limiar, Olson (2001, p. 659) propõe “[...] a adoção de uma postura ativa criando-se espaços em nossos limites para as vozes dos que foram excluídos”. Assim, metodologias, técnicas e instrumentos que façam com que os sistemas de informação sejam mais acessíveis e colaborativos podem corroborar para o desenvolvimento de relações éticas nesse contexto.

Essa proposição pode ser justificada pelo fato de que, atualmente os usuários ainda se “deparam com dispositivos coercitivos na prática bibliotecária [ou de profissionais da informação] que limitam o acesso e uso da informação”, os quais suprimem a verdadeira relação do usuário com o conhecimento (Aquino; Santana, 2013). Esses dispositivos, por sua vez, estão vinculados às práticas de organização e

representação do conhecimento e, em determinados casos, podem implicar negativamente na comunicação entre a linguagem dos usuários e a linguagem dos sistemas de recuperação da informação.

Santos *et al* (2017) enfatizam que essas reflexões se mostram imprescindíveis às perspectivas de formação e do trabalho do indexador, considerando a necessidade de buscar uma linguagem que não seja excludente e questione modelos de classificação cuja inflexibilidade impede que se represente a diversidade de maneiras menos díspares.

Diante do exposto, é possível refletir que as garantias citadas anteriormente, em conjunto com a hospitalidade cultural, podem fornecer a base para a análise, seleção e atualização de termos; bem como para o estabelecimento de relações semânticas entre eles de forma apropriada em sistemas híbridos de Organização do Conhecimento, considerando o conjunto de crenças, valores, atividades e históricas representadas e compartilhadas colaborativamente pelas pessoas em um dado domínio do conhecimento.

Em analogia a essas reflexões, é importante ressaltar a necessidade da inclusão de abordagens e metodologias colaborativas e integradoras que considerem os termos que representem efetivamente a linguagem, a diversidade e a cultura dos usuários em processos, instrumentos, produtos e serviços voltados para a organização, tratamento, representação, recuperação, recomendação, acesso e uso da informação. Isso se deve ao fato de que a linguagem é um dos meios pelos quais as ideias são representadas para e por uma cultura.

Portanto, acredita-se que as garantias citadas anteriormente podem corroborar para a produção significativa de metadados semânticos, bem como para a noção de diversidade nas práticas de representação do conhecimento.

3 FOLKSONOMIAS E REPRESENTAÇÃO COLABORATIVA DA INFORMAÇÃO

A Folksonomia faz parte de uma terminologia cunhada em 2004 por Thomas Vander Wal com a junção das palavras “*Folk*” (do germânico: “povo”, “grupo de pessoas”) e “*taxonomy*” (do grego: “ciência ou técnica de classificação”). De acordo com Quintarelli (2005), o termo em si nasceu na busca de nomear uma atividade já implementada em ferramentas de sites como o Delicious e o Flickr.

Para Santos e Corrêa (2018) as Folksonomias podem ser conceituadas como o resultado da atribuição livre de palavras-chave realizada por usuários (humanos ou robôs), sem o auxílio de instrumentos de controle terminológico, em sistemas colaborativos para fins de representação e recuperação de objetos informacionais em diferentes formatos, sejam eles áudios, textos, imagens, vídeos etc.

O interesse pelas investigações sobre esse tema despontou em escolas de informação por meio de ensino e pesquisa com ênfase tecnológica, principalmente em institutos e faculdades de Ciência da Informação alemães e norte-americanos.

As falhas dos sistemas que utilizam as Folksonomias decorrem, em linhas gerais, dos erros ortográficos, das *tags*

excessivamente personalizadas, da ambiguidade terminológica em etiquetas que representam os materiais informacionais e, mais recentemente, do uso de *chatbots* programados para atribuírem *tags* que propagam a rápida popularização de *fake news* (notícias falsas) na rede. Por outro lado, se bem aplicada e mediada, as Folksonomias podem trazer significativas contribuições para as práticas de representação e recuperação de conteúdos em ambientes digitais, dentre as quais: auxílio no fornecimento de detalhes nas descrições dos objetos informacionais (imagens, textos, áudios, vídeos, etc.); apoio na construção e atualização de instrumentos de controle terminológico para a representação e recuperação da informação.

Na concepção de Mika (2005), as *tags* se configuram como entidades que referenciam a representação colaborativa do conhecimento entre os atores sociais, mediante o compartilhamento de narrativas, linguagens e signos comuns.

Um das principais diferenças entre as Folksonomias e os vocabulários consiste na inexistência de relações hierárquicas entre os termos, podendo os mesmos serem apresentados em conjunto no âmbito de um sistema colaborativo. Dessa maneira, as Folksonomias consistem em um sistema colaborativo de organização do conhecimento que apresenta fragilidade em relação ao grau de formalização da linguagem, mas que possui semântica implícita contribuindo para a encontrabilidade e serendipidade em um domínio de especialidade.

Segundo Mathes (2004), as Folksonomias contribuem para a serendipidade, ou seja, a descoberta casual de

informações em ambientes digitais, a partir dos *links* gerados entre as *tags*, os quais contribuem para o processo de navegação em ambientes colaborativos.

Nesse limiar, as Folksonomias são compreendidas como um sistema democrático, descentralizado e inclusivo, uma vez que agregam e consideram o vocabulário de todas as pessoas, de todas as vozes e de diferentes pontos de vista, viabilizando variados pontos de acesso às informações (Quintarelli, 2005). Nesse entendimento, Golder e Huberman (2005) ressaltam que as mesmas podem ser vistas como um sistema alternativo e mais leve de classificação, uma vez que a finalidade de uma *tag* é organizar os registros do conhecimento representados pelos autores desses recursos.

A pesquisa realizada por Adler (2009) revela duas características importantes da representação colaborativa da informação: a) a existência de um grau de consenso por parte dos usuários quanto à etiquetagem dos recursos; b) alcance visível e negociável das expressões de vozes minoritárias nos ambientes colaborativos, uma vez que as *tags* atribuídas são estabelecidas pelas convenções sociais da linguagem e da ideologia dos usuários.

Desse modo, concorda-se com Santos *et al* (2017), quando enfatizam que as Folksonomias refletem as relações entre a linguagem e a comunidade, resultando em práticas e vocabulários diferentes daqueles presentes em discursos autorizados e normalizados (como os tesouros, as taxonomias e as ontologias, por exemplo).

Em consonância de evidências, Yedid (2013) evidencia que as Folksonomias refletem o vocabulário da comunidade de usuários e contribuem para a identificação de tendências terminológicas entre atores sociais que descrevem um *corpus* de documentos ou itens específicos, podendo ser analisadas para selecionar termos preferidos dessas fontes, bem como criar ou atualizar instrumentos de controle terminológico.

Spiteri (2007) realizou um estudo em três sites (delicious, Furl e Technorati), na oportunidade constatou que as etiquetas analisadas nesses ambientes encontravam-se articuladas com uma quantidade significativa de orientações da National Information Standards Organization (NISO) para a construção, desenvolvimento e gerenciamento de vocabulários controlados – sobretudo em relação aos tipos de conceitos expressos, a predominância das *tags* e substantivos únicos, bem como o uso de ortografia reconhecida.

Como mencionado anteriormente, a sinergia das Folksonomias e dos instrumentos de controle terminológico pode apresentar benefícios para os sistemas colaborativos de representação e recuperação da informação na web semântica. Essa cooperação pode ser utilizada na construção de sistemas híbridos de organização do conhecimento, a fim de construir colaborativamente representações do conhecimento com base na linguagem da comunidade de usuários de domínios específicos, considerando os preceitos estabelecidos nas garantias apresentadas nesta pesquisa.

3.1 Proposições de Sistemas híbridos de Organização do Conhecimento

Dada a versatilidade e capacidade da Folksonomia para integrar outros vocabulários, as propostas existentes na literatura nacional e internacional evidenciam a necessidade de conciliar as bases folksonômicas com as estruturas de instrumentos de controle terminológico (como taxonomia e ontologia). As proposições, em sua maioria, objetivam apoiar as interações sociais e auxiliar o ciclo de vida das Folksonomias e ontologias, enquanto sistemas que sustentam as bases de conhecimento das aplicações web.

Mika (2005), por exemplo, realiza a extração de relações mais amplas e estreitas entre as *tags* para o desenvolvimento de ontologias leves através de um conjunto de termos conectados a partir de relacionamentos semânticos limitados.

Por outro lado, as Folksonomias também podem ser enriquecidas semanticamente por outros Sistemas de Organização do Conhecimento com base na estrutura sistematizada das ontologias ou nas ferramentas da web semântica para auxiliar na representação e navegação, bem como sugerir aos usuários os termos semanticamente relacionados às suas consultas.

Seguindo essa lógica, Gruber (2005) apresenta uma proposta denominada “TagOntology” cujo objetivo é desenvolver uma ontologia dedicada a formalizar a ação de etiquetagem a partir dos seguintes elementos: o objeto (item informacional etiquetado); as etiquetas utilizadas na

marcação; o usuário que realizou a etiquetagem; e o domínio no qual a etiquetagem é realizada.

O primeiro passo registrado na literatura por pesquisadores para processar dados extraídos das Folksonomias para a construção de instrumentos de controle terminológico tem sido calcular a similaridade de *tags* e inferir as relações semânticas entre etiquetas. Specia e Motta (2007), por exemplo, desenvolveram um sistema que aplica uma sequência de tratamentos semânticos às Folksonomias, dentre eles: busca de *tags* equivalentes; agrupamento de etiquetas semelhantes com base em cálculos de medidas de similaridade de acordo com as etiquetas. Para tanto, são empregadas estratégias de pré-processamento superficiais e técnicas estatísticas, juntamente com ferramentas ontológicas fornecidas pela web semântica.

Os resultados das análises realizadas por esses autores geraram *clusters* com *tags* altamente relacionadas e correspondentes a conceitos em ontologias, apresentando relacionamentos semânticos significativos.

As taxonomias são Sistemas de Organização do Conhecimento amplamente aplicados em estruturas de navegação de sites e diretórios para a recuperação de conteúdos. Contudo, essas ferramentas têm apresentado limitações no que concerne às nuances culturais e nomenclaturas utilizadas em sua configuração. Com base nessa perspectiva, autores como Kiu e Tsui (2010) constataam que as Folksonomias e taxonomias, em conjunto, contribuem para a melhoria da navegação e personalização da busca e recuperação do conhecimento. Para tanto, propõem um

algoritmo que integra uma classificação híbrida com a taxonomia e a Folksonomia utilizando métodos heurísticos de mineração de dados, objetivando o aprimoramento na navegação do conhecimento.

Portanto, depreende-se que algumas propostas na literatura se direcionam para a criação de instrumentos de controle terminológico a partir dos dados gerados pelas Folksonomias. Por outro lado, algumas abordagens conectam as Folksonomias a instrumentos de controle terminológico pré-existentes. Em ambos os casos é possível identificar uma unidirecionalidade, em que um sistema enriquece o outro. Seguindo essa linha de raciocínio, a proposta de Alves (2012) contempla uma perspectiva bidirecional, em que ontologias e Folksonomias são combinadas em uma nova entidade denominada “Ontologia folksonomizada”, harmonizando aspectos complementares de ambas.

Além das melhorias constatadas nas práticas de indexação e recuperação da informação, o autor supracitado propôs uma técnica chamada *3E Steps* com vistas a prestar suporte ao desenvolvimento de ontologias a partir de dados das Folksonomias. Também foi implementado um protótipo de uma ferramenta que objetiva a construção de ontologias folksonomizadas e suporte à revisão de ontologias.

Em linhas gerais, observa-se que as propostas anteriormente descritas são relevantes para o desenvolvimento de ambientes colaborativos, uma vez que os seus métodos ou algoritmos contribuem para a criação de ferramentas mais precisas cujo objetivo é melhorar as experiências de navegação, representação, busca e

recuperação semântica dos registros do conhecimento. Essas propostas podem ser integradas a modelos colaborativos de indexação, os quais de acordo com Santos (2016) são modelos que delineiam atividades colaborativas de indexação a serem executadas por usuários no processo de representação colaborativa da informação e servem como referências para o desenvolvimento de sistemas colaborativos, podendo ser adaptados ou aperfeiçoados em diversos contextos.

4 CONSIDERAÇÕES FINAIS

Face ao exposto, percebe-se que a representação no contexto da Ciência da Informação conquista espaço não apenas como uma operação pragmática, mas se amplia como campo que abrange estudos teóricos e aplicados que objetivam tornar o uso e a apropriação do conhecimento produzido, devendo considerar questões éticas, semânticas, organizacionais e socioculturais no intuito de tornar os conteúdos disponíveis e baseados na linguagem dos seus usuários.

As operações de indexação, construção e atualização de instrumentos de controle terminológico tem influências de fatores linguísticos, cognitivos, lógicos, ideológicos, contextuais, culturais, sociais inerentes ao profissional da informação, além do fato de que o olhar leitor desse profissional pode silenciar ou expor diferentes leituras.

Sujeitos que fazem parte de determinadas culturas podem não conseguir (ou apresentar dificuldades para) reconhecer, assimilar, interpretar termos específicos pelo fato

de não possuírem conhecimentos adequados para entendê-los. Tal fato evidencia a importância de considerar as novas possibilidades de representação colaborativa da informação para a produção significativa de metadados semânticos em sistemas.

Diante da capacidade e versatilidade das Folksonomias para integrar outros Sistemas de Organização do Conhecimento, bem como da possibilidade de ser enriquecida semanticamente por instrumentos como ontologias, foi possível visualizar propostas de Sistemas híbridos de Organização do Conhecimento. Esses são conceituados nesta pesquisa como sistemas que agregam as potencialidades das Folksonomias e de outros instrumentos (ontologia, taxonomia, tesouro, por exemplo), podendo ser aplicados em domínios específicos, objetivando a produção significativa de metadados semânticos para fins de representação e recuperação de conteúdos na web.

Essas propostas exploram as potencialidades da Folksonomia com o suporte dos formalismos e métodos utilizados na construção de instrumentos de controle terminológicos, descortinando novas perspectivas para auxiliar a troca de conhecimentos em ambiente web.

É importante ressaltar que a interoperabilidade deve ser considerada no futuro das aplicações que integram as Folksonomias, sendo necessária a criação de modelos para adequar os sistemas colaborativos aos usos, tecnologias e padrões ancorados nos preceitos das garantias e da web semântica, sobretudo na perspectiva dos dados abertos interligados. Tais mudanças consistem em desafios, uma vez

que objetivam alcançar uma combinação apropriada de precisão, diversidade, semântica e representatividade no que concerne às representações dos conhecimentos de comunidades discursivas em cenários colaborativos.

Em linhas gerais, depreende-se que a capacidade de lidar com percepções plurais é algo que deve constar na lista de aptidões almejadas por pesquisadores e profissionais da informação dispostos a trabalhar com as atividades de construção e atualização de Sistemas de Organização do Conhecimento, bem como de ferramentas, processos, sistemas e serviços baseados nas ferramentas tecnológicas contemporâneas.

Além disso, considera-se importante que os pesquisadores e profissionais da informação dediquem-se aos estudos de Inteligência Artificial e dos ambientes algoritmicamente modelados com vistas a contribuir para a produção significativa de metadados semânticos e a recuperação de diferentes objetos informacionais (textos, áudios, imagens, vídeos etc.).

REFERÊNCIAS

ADLER, M. Transcending Library Catalogs: A comparative study of Controlled terms in Library of Congress Subjects Headings and User-Generated Tags in Libraything for transgender Books. *Journal of web librarianship*, [United States], v. 3, n. 1, p. 309-331, 2009.

ALVES, H. A. *Ontologias folksonomizadas: uma abordagem para a fusão de ontologias e folksonomias*. 2012. Dissertação

(Mestrado em Ciências da Computação) – Universidade Estadual de Campinas, Campinas, 2012.

AQUINO, M. A.; SANTANA, V. A. Práticas de organização e representação da informação étnico-racial em bibliotecas universitárias: necessidade de preservação da memória de negros. *RICI: Revista Ibero-americana de Ciência da Informação*, Brasília, v. 6, n. 2, p. 17-36, ago./dez. 2013.

BARITÉ, M. La garantía cultural como justificación em sistemas de organización del conocimiento: aproximación crítica. *Palabra Clave*, La Plata, v. 1, n. 1, p. 2-11, 2011.

BARITÉ, M. La garantía literária: vigência y proyección teórico-metodológica. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 13., Salvador, 2007. *Anais eletrônicos [...]*. Salvador: ANCIB, 2007.

BEGHTOL, C. Bibliographic Classification Theory and Text Longuistics: Aboutness Analysis, Intertextuality and the Cognitive Act of Classifying. *Journal of Documentation*, Bradford, v. 42, n. 2, p. 84-113, 1986.

BEGHTOL, C. A proposed ethical warrant for global knowledge representation and organization systems. *Journal of Documentation*, London, v. 58, n. 5, p. 507-532, 2002.

BEGHTOL, C. Ethical decision-making for knowledge representation and organization systems for global use. *Journal of the American Society for Information Science and Technology*, New York, v. 56, n. 9, p. 903-912, 2005.

FRANCELIN, M. M.; PINHO, F. A. *Conceitos na Organização do Conhecimento*. Recife: Ed. Universitária da UFPE, 2011.

GOLDER, S.; HUBERMAN, B. A. *The structure of collaborative tagging systems*. 2005. Disponível em: https://www.researchgate.net/publication/1958429_The_Str

ecture_of_Collaborative_Tagging_Systems. Acesso em: 02 maio 2021.

GRUBER, T. Ontology of Folksonomy: A Mash-up of Apples and Oranges. *In: CONFERENCE ON METADATA AND SEMANTICS RESEARCH (MTSR), [S. l.], 2005. Proceedings [...]. [S. l.: s. n.]: 2005.*

GUEDES, R. M. *O princípio da garantia semântica e os estudos da linguagem*. 2016. Tese (Doutorado em Ciência da Informação) – Universidade Federal de Minas Gerais, Belo Horizonte, 2016.

HALL, S. *Cultura e Representação*. Rio de Janeiro: Ed. PUC-Rio: Apicuri, 2016.

KIU, C.; TSUI, E. TaxoFolk: A hybrid taxonomy-folksonomy classification for enhanced knowledge navigation. *Knowledge Management Research & Practice*, [S. l.], v. 8, n. 1, p. 24-32, 2010.

KWASNIK, B. H. Semantic Warrant: a pivotal concept for our field. *Knowledge Organization*, Baden-Baden, v. 37, n. 2, p. 106-110, Apr. 2010.

LANCASTER, F. W. *Indexação e resumos: teoria e prática*. 2. ed. Brasília: Briquet de Lemos / Livros, 2004.

LIMPENS, F.; GANDON, F.; BUFFA, M. *Linking folksonomies and ontologies for supporting knowledge sharing: a state of the art*. [S. l.: s. n.], 2009.

MATHES, A. *Folksonomies: cooperative Classification and Communication Through Shared Metadata*. 2004. Disponível em: <https://adammathes.com/academic/computer-mediated-communication/folksonomies.html>. Acesso em: 20 maio 2021.

MIKA, P. Ontologies are Us: a Unified Model of Social Networks and Semantics. *In: INTERNATIONAL SEMANTIC WEB CONFERENCE, 4., 2005, Galway. Proceedings [...]. [S. l.]: Springer, 2005. p. 522-536.*

OLSON, H. A. The power to name: representation in library catalogs. *Journal of Women in Culture and Society, [S. l.], v. 26, n. 1, p. 639-68, 2001.*

QUINTARELLI, E. Folksonomies: power to the people. *In: ISKO ITALY-UNIMIB MEETING, 2005, Milan. Proceedings [...], Milán: ISKO, 2005. Disponível em: <http://www.iskoi.org/doc/folksonomies.htm>. Acesso em: 20 maio 2021.*

SANTOS, R. F. *Modelos colaborativos de indexação e a sua aplicabilidade na Base de Dados Referencial de Artigos de Periódicos em Ciência da Informação*. 2016. Dissertação (Mestrado em Ciência da Informação) – Universidade Federal de Pernambuco, Recife, 2016.

SANTOS, R. F. *et al.* A representação colaborativa da informação e a construção de linguagens documentárias sobre diversidade de gêneros: análise das contribuições do Dicionário de Gêneros – “só quem sente pode definir”. *In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 18., 2017, Marília. Anais eletrônicos [...]. Marília: ANCIB, 2017.*

SANTOS, R. F.; CORRÊA, R. F. Análise das definições de Folksonomia: em busca de uma síntese. *Perspectivas em Ciência da Informação*, Belo Horizonte, v. 23, n. 2, p. 1-32, abr./jun. 2018.

SOLER MONREAL, C.; GIL LEIVA, I. Posibilidades y límites de los tesauros frente a otros sistemas de organización del conocimiento: folksonomías, taxonomías y ontologías. *Revista*

Interamericana de Bibliotecología, Medellín, v. 33, n. 2, p. 361-377, jul./dic. 2010.

SPECIA, L.; MOTTA, E. Integrating folksonomies with the semantic web. *In: EUROPEAN CONFERENCE ON THE SEMANTIC WEB*, 4., 2007, Innsbruck. *Proceedings* [...]. Berlin: Springer-Verlag, 2007.

SPITERI, L. F. Structure and form of folksonomy tags. The road to the public library catalogue. *Webology*, [S. l.], v. 4, n. 2, 2007.

YEDID, N. Introducción a las Folksonomías: definición, características y diferencias con los modelos tradicionales de indización. *Información, cultura y sociedad*, [S. l.], v. 1, n. 29, p. 13-26, 2013.

ZAMBONI, R. C. V. *Organização do conhecimento, classificação e diversidade cultural: uma análise a partir do conceito de "garantias"*. 2018. Tese (Doutorado em Ciência da Informação) – Escola de Comunicações e Artes, Universidade de São Paulo, São Paulo, 2018.

CAPÍTULO 7

RESGATE DOS FUNDAMENTOS DA CATALOGAÇÃO DE ASSUNTO: PRINCÍPIOS, ENTRADAS PADRONIZADAS E REMISSIVAS EM FOCO NA ANÁLISE DO SITE DA *AMAZON*

Brisa Pozzi de Sousa

1 DELINEAMENTO INICIAL: O RESGATE DA CATALOGAÇÃO DE ASSUNTO

O foco desta reflexão é discorrer sobre alguns aspectos da representação de assunto, pela catalogação de assunto (enquanto processo), especificamente na sua função analógica e em comparação ao contexto digital. Para isso, a escolha foi o site da *Amazon* (amazon.com.br), que vinculado a empresa norte-americana de mesmo nome tem faturamento e crescimento considerável (Farani, 2020), sobretudo no atual cenário pandêmico. Ademais, o cenário impactado pela covid-19 potencializou o comércio eletrônico, pois ficar em casa e permanecer em isolamento social é uma forma das pessoas se protegerem.

Não se pode negar a crescente busca pelo comércio eletrônico, mas analisar convergência(s) em aspectos de formas de representação, e que há bastante tempo faz parte da Biblioteconomia, constitui o objetivo desse trabalho.

Especificamente, a representação de assunto relaciona-se a diferentes perspectivas e recortes geográficos que se conectam aos fundamentos teóricos, permanecendo a compreensão da função de recuperação da informação. Objetivando indicar que a Catalogação de Assunto (com letra maiúscula para diferenciar a área do processo) é atuante na Biblioteconomia, mas também pode ser resgatada para além dos catálogos de biblioteca, busca-se revisitá-la para canalizar sua proposta de princípios, entradas padronizadas e remissivas.

Considera-se oportuno que o resgate de fundamentos relativos à Catalogação de Assunto possa ser divulgado, quer seja para a consolidação de pesquisas teórico-prática, quer seja como disciplina (ou parte) da formação acadêmica de estudantes de Biblioteconomia. Interessante seria na mesma medida o ambiente de pós-graduação em Ciência da Informação fornecer promoção as disciplinas que aparentam ser obsoletas, mas que na verdade contribuem para a consolidação científica da organização da informação.

Assim, traçar um breve itinerário percorrendo alguns tópicos da Catalogação de Assunto, para apontar origens remotas de ideias aparentemente novas que têm guiado os caminhos da Ciência da Informação, especificamente nos moldes do que é considerado tecnologia (e tecnologia não apenas e somente no emprego de computadores), é a motivação dessa análise. Reduzir a diversidade dos fundamentos e focar somente na aplicação prática reflete pobreza conceitual. Portanto, as reflexões aqui expostas partem de uma forte motivação de resgate para aprofundar e

clarificar ideias, num ambiente de proveitoso diálogo, que não exige críticas e discordância dos pares.

2 A SISTEMATIZAÇÃO DOS PRINCÍPIOS, ENTRADAS PADRONIZADAS E REMISSIVAS POR CHARLES CUTTER NA CATALOGAÇÃO DE ASSUNTO

Charles Ami Cutter (1837-1903) é considerado o grande teórico da Catalogação de Assunto (*Subject Cataloguing*) na literatura de origem anglo-saxã. A organização de catálogos por assunto nos Estados Unidos foi tratado por Cutter, criador de regras que culminaram com a difusão do catálogo dicionário (Taylor, 1948). Buonocore (1976, p. 118, tradução nossa) define esse catálogo como a fusão de três diferentes tipos: de autor, de assunto e de título. Para este autor, a dificuldade a ser enfrentada em relação ao catálogo dicionário, é a “[...] escolha apropriada da palavra-chave [...] que sirva para caracterizar o assunto tratado na obra, escolha que o catalogador faz com o auxílio de listas [...] para tais fins.”¹

Mann (1962, p. 151) explica que no catálogo dicionário “[...] tôdas [sic] as palavras de entrada (autor, título, assunto) são arrumadas numa só lista alfabética. [...] Esse catálogo é assim chamado por sua arrumação semelhante a um dicionário.” Ainda segundo a mesma autora, palavra de entrada é o mesmo que palavra-chave, responsável em ocupar a parte da entrada sob a qual a ficha é arquivada no catálogo.

¹ “[...] acertada elección de la palabra clave [...] que sirve para caracterizar el asunto de que trata la obra, elección que hace el catalogador con el auxilio de listas [...] para tales fines.” (Buonocore, 1976, p. 118).

Logo, a ficha, que é o suporte do registro de um documento, leva em consideração diferentes atributos para sua entrada.

Cutter trabalhou com Charles Jewett na *Public Library of Boston*, objetivando a construção de um catálogo para uma coleção especial (Fiúza, 1987). Todavia, quando Cutter ainda era aluno da *Harvard Divinity School*, tornou-se auxiliar de biblioteca, entre 1857-1859, e o cargo foi determinante para sua carreira na criação da sua obra mais conhecida, as *Rules for a Printed Dictionary Catalogue*. O então bibliotecário e catalogador chefe, Ezra Abbot, notou seu desempenho e, em 1860, convidou-o para a equipe do *Harvard College*. Já no final do ano de 1868, em decorrência do trabalho desenvolvido com Abbot, Cutter recebeu outro convite: ser bibliotecário do *Boston Athenaeum* que, junto com *Harvard*, eram os pontos principais da influente comunidade intelectual de Boston-Cambridge (Miksa, 2012; World Encyclopedia of Library and Information Services, 1993). Ao assumir o cargo no *Boston Athenaeum*, ele destacou como primordial “[...] a compilação de um catálogo impresso. Usou, então, a sua experiência profissional influenciada diretamente por Jewett e, indiretamente por Panizzi, [...] para estabelecer suas <<Regras para um Catálogo Dicionário Impresso>>.” (Fiúza, 1987, p. 48).

Garrido Arilla (1999, p. 80, tradução nossa) aponta que as *Rules* constituíram um código que permitia “[...] elaborar um registro bibliográfico por três entradas: autor, título e assunto.”² Anteriormente, no entanto, Anthony Panizzi estruturou nas 91 Regras orientações para a criação de

² “[...] redactar un registro bibliográfico por tres entradas: autor, título y materia.” (Garrido Arilla, 1999, p. 80).

catálogos alfabéticos, principalmente de autores pessoais e corporativos, que permitissem ao público a localização rápida e fácil de uma obra, além do agrupamento das obras de um mesmo autor com suas distintas edições e traduções.

Mas, Cutter, ao desenvolver as *Rules for a Printed Dictionary Catalogue* (1876), que ficou conhecida como *Rules for a Dictionary Catalog*³, preocupou-se em construir regras (369 ao total) que norteassem a decisão das entradas no catálogo, incluindo a de assunto (Garrido Arilla, 1999). Sem tais orientações, o catalogador adotaria diferentes cabeçalhos para os mesmos assuntos, gerando dúvidas no processo de busca pelo público. Dentre as regras de Cutter, muitas delas têm relação direta com a entrada no catálogo.

Segundo Mann (1962, p. 176) “por entrada de assunto queremos dizer o termo ou termos usados num catálogo [...] para exprimir o assunto ou assuntos dos livros.” De acordo com essa autora, as *Rules* tornaram-se a principal referência, em inglês, para o estabelecimento de regras na elaboração de cabeçalhos de assunto, que formaram as listas. Além de termos, palavras, frases, nomes ou expressões foram empregadas para reunir as listas de cabeçalhos de assunto. Ademais, Cutter (1904, p. 21) também definiu cabeçalho (*heading*) como a palavra por ordem alfabética que fornece o

³ As regras para catálogos do tipo dicionário de Cutter receberam o título *Rules for a Printed Dictionary Catalogue* nas primeiras edições: 1ª (1876), 2ª (1889) e 3ª (1889). Já a 4ª e última edição, de 1904, foi intitulada *Rules for a Dictionary Catalog* (World Encyclopedia of Library and Information Services, 1993).

lugar da entrada no catálogo e pode ser determinado pelo nome do autor, do assunto ou palavra do título.

Em referência ao cenário retratado até aqui, nota-se que elaborar orientações para as entradas dos catálogos não foi uma preocupação apenas de Cutter, mas ele criou orientações normativas para dar forma à representação alfabética no catálogo e, assim, por mais de meio século as *Rules* permaneceram como a única sistematização disponível para pensar e estruturar os assuntos pelos cabeçalhos.

Com o acesso às bibliotecas pela população⁴, constata-se a continuidade de alteração da percepção do catálogo, de simples ferramenta de registro e inventário do acervo, para instrumento propiciador do livre acesso às consultas das bibliotecas, inclusive as públicas. Conforme demonstrado, tais mudanças decorreram ao longo dos séculos sob variadas influências até alcançar orientações consistentes para a organização dos catálogos. No entanto, o trabalho é tomado como útil quando indivíduos incorporam essa organização para alterar o seu estado de conhecimento. É nesse ponto que a demanda por informação organizada, necessitou de parâmetros para pensar os assuntos dos documentos e se,

⁴ A 'população' que fazia uso das bibliotecas era constituída por um público restrito de pessoas, ou seja, um público singular assíduo a esse ambiente com uma maneira própria de buscar informação. Cutter (1904) usou o termo '*the users*'. Ao utilizar o artigo *the* antes do substantivo, o autor dá indícios de considerar a existência de outros grupos de usuários, mas presume-se a homogeneidade de um público à época: homem branco, heterossexual e com certo nível de escolaridade. Para Olson (2001, p. 641, tradução nossa), a lógica utilizada por Cutter "[...] introduz uma aparente imposição democrática de que o catálogo deveria ser construído para a conveniência do público ao qual ele serve".

anteriormente, os próprios bibliotecários atuavam como catálogos vivos, para auxiliar o acesso aos acervos, tal situação se tornou inconcebível (Shera; Egan, 1969).

Cutter recebeu apoio do *Bureau of Education*, órgão do governo dos Estados Unidos, que por sua vez, solicitou um relatório sobre a história, condição e administração das necessidades das bibliotecas públicas. Com isso, as *Rules* foram concluídas em 1875, e publicadas em primeira edição como a parte dois (*Part II*) do relatório *Public Libraries in the United States of America: their history, condition and management* (Miksa, 2012; World Encyclopedia of Library and Information Services, 1993).

Essa divulgação sinalizou, mesmo que indiretamente, a importância de um instrumento sistematizador, aprovado por um órgão do governo, que desse conta de indicar os atributos dos documentos e a forma de organização na ficha. Os acervos gerais das bibliotecas públicas demandaram uma organização para o seu uso.

Mas, em que medida o assunto foi uma preocupação nas *Rules*, visto ter sido um dos documentos mais conhecido em língua inglesa e responsável por indicar essa sistematização? Para essa compreensão, primeiro é necessário observar à abordagem das *Rules*, pois sua divisão preponderou em duas grandes partes, sendo a primeira (ou A) dedicada à Entrada (*A. Entry – Where to enter*, ou, Onde entrar) e, a segunda (ou B), dedicada ao Estilo (*B. Style – How to enter*, ou, Como entrar). Antecedendo as duas, encontram-se os Comentários Gerais (*General Remarks*), que se subdivide em

Objetivos (*Objects*), Meios (*Means*) e Razões para escolha (*Reasons for choice*).

Formas para a 'Entrada' no catálogo já haviam sido previstas por Panizzi, mas o 'Estilo' toma significância por propor critérios pragmáticos para uma política do catálogo. Toma-se proximidade à Entrada e a busca no catálogo. Cutter propõe os Objetivos do catálogo, como segue (Cutter, 1904, p. 12, tradução nossa):

1. Permitir a uma pessoa encontrar um livro do qual
(A) o autor
(B) o título
(C) o assunto } **é conhecido**
2. Mostrar o que a biblioteca possui
(D) por um autor determinado
(E) com um assunto determinado
(F) em um tipo de literatura determinado
3. Ajudar na escolha de um livro
(G) quanto a sua edição (bibliograficamente)
(H) quanto ao seu caráter (literário ou tópico).⁵

O primeiro item propõe permitir o encontro de um livro através de ao menos um dos três diferentes elementos: autor-

⁵ 1. To enable a person to find a book of which either
(A) the author
(B) the title } is known
(C) the subject

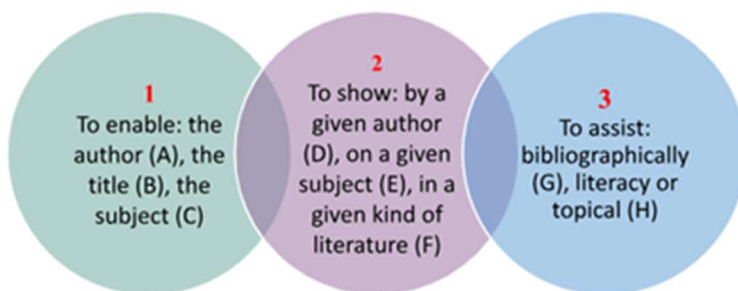
2. To show what the library has
(D) by a given author
(E) on a given subject
(F) in a given kind of literature

3. To assist in the choice of a book
(G) as to its edition (bibliographically)
(H) as to its character (literacy or topical). (Cutter, 1904, p. 12).

A, título-B e assunto-C. O segundo item busca mostrar o que a biblioteca possui, por ao menos um dos três elementos: autor-D, assunto-E e tipo de literatura-F, sendo que autor-D e assunto-E, também estão inseridos na possibilidade anterior (A e C). Já o terceiro item consiste em auxiliar a escolha de um livro por dois elementos: bibliograficamente-G e literário ou tópico-H. O esquema apresenta os elementos necessários para buscas no catálogo e isentam relações menos habituais à época.

De outro modo, no primeiro objetivo, pela identificação de um livro no catálogo é possível alcançar o segundo objetivo, que mostra o que a biblioteca tem, possibilitando o agrupamento de atributos dos livros. Já o terceiro objetivo ajuda a escolher um livro por sua edição, ou caráter literário ou tópico, pois o usuário ao encontrar livros similares necessitará realizar uma seleção, por exemplo, entre as várias edições disponíveis. É possível refletir, ainda, que no século XIX não havia abundância de muitas edições dos livros, mas foi ponto merecedor de atenção.

Figura 1: Sistematização dos três Objetivos das *Rules* de Cutter (1904)



Fonte: A autora.

A interseção apresentada na figura 1 é explícita em A e D, C e, mas não da mesma maneira em F e H. Fiúza (1987, p. 49), explica que para alcançar os Objetivos, que ela denominou de “Objetos”, o catálogo conta com seis Meios: “1. Entradas de autor, com as referências necessárias (A e D); 2. Entradas de título ou referências ao título (B); 3. Entradas de assunto, referências cruzadas (C e); 4. Entrada de forma e língua (F); 5. Transcrição da edição e imprensa, com notas, quando necessário (G); 6. Notas especiais (H).” Em alguma medida, os Objetivos aproximam-se dos Meios, e os aspectos principais são estabelecidos pelos atributos autor, título e assunto que moldam a forma da entrada do catálogo dicionário. O assunto propiciaria a referência cruzada (*cross reference*), que remete um assunto para outro, e essa recomendação de Cutter consistiu na criação de uma rede de remissivas, pelo Princípio Sindético, para os catálogos. Mas, a preocupação central não foi com o assunto e sim com a sua forma de entrada.

Se a rede era limitada por questões econômicas, a preocupação com a fundamentação do assunto ficou em segundo plano e não houve a difusão de algo diferente do que já era realizado. Nesse sentido e conforme Strout (1956), desde o início do século XVIII as referências cruzadas já eram bastante comuns. A complexidade das técnicas de organização cresceu “à medida que aumentavam os impressos tanto em número como em variedade [...]” (Shera; Egan, 1969, p. 11).

Buonocore (1976, p. 367) explica que o termo referência (*referencia*), no sentido amplo, está relacionado com o processo de catalogação nas chamadas referências cruzadas (*referencias cruzadas*) “ver” (*véase*) e “ver também”

(véase también). Portanto, há que se considerar nesse contexto, remissiva e referência como sendo palavras correlatas.

Ainda em relação aos elementos F, G e H dos Objetivos, um tipo de literatura determinado, tanto à edição (bibliograficamente) quanto o caráter (literário ou tópico) do livro poderiam compor uma outra forma de reunião, pois os livros que detém várias edições poderiam ser representados não como uma entidade separada, mas sim de forma agrupada, situação que comungaria com o Princípio de Uso.

Pautando-se em Garrido Arilla (1999), segue a descrição dos três princípios de Cutter: 1) Princípio Específico: o livro deve ser incluído no catálogo com o cabeçalho de assunto mais específico possível, fator esse determinante para a economia e uniformidade do instrumento; 2) Princípio de Uso: os cabeçalhos de assunto devem utilizar termos que os usuários do catálogo esperem encontrar; 3) Princípio Sindético: os cabeçalhos de assunto empregados nos catálogos devem ter conexão com referências que conduzam o usuário de uma forma que não é autorizada ou utilizada, para a que seja autorizada e utilizada pelos usuários. Verifica-se nas *Rules* que os Princípios em sua ordem são guiados, respectivamente, pelos números §161, §169 e §187 (Cutter, 1904).

Ao tratar das Razões para escolha (*Reasons for choice*), a fim de alcançar os Objetivos propostos, Cutter (1904) explica que pode haver o emprego de vários métodos. Entretanto, ele relaciona três pontos e os vincula à forma de entrada no catálogo dicionário. Primeiro, indica a necessidade de observar o público de pessoas que usam a biblioteca; segundo, a

consistência e cobertura das entradas; e, por último, se for uma entrada ampla para um grande público, que ele chamou de entradas de massa (*mass entries*), deve-se perceber as possíveis dificuldades em empregar essa entrada e adotar a que melhor auxilie a busca. Os três tópicos são bastante abrangentes, não especificam nenhum método, não aludem aos elementos dos Objetivos, mas se aproximam dos três Princípios.

Cutter desenvolveu seus fundamentos com foco no catálogo manual, compreendido hoje como analógico. Nos catálogos manuais, pontos de acesso são criados e estruturados para facilitar o conhecimento do acervo da biblioteca e, também, a recuperação da informação. Nesse sentido, o catálogo é construído com foco na elaboração de índices alfabéticos para que usuários realizem buscas (com base nas suas necessidades individuais de informação) e recuperam documentos por seus autores, pelo título do documento ou da coleção, ou por assuntos. Nos catálogos online (que são digitais) há ampliação dos pontos de acesso e, por consequência, ampliam-se também as formas de recuperação.

Como modo de comparação e, tomando por base um catálogo de produtos impresso (que é analógico) e um catálogo de produtos online (digital), na segunda modalidade ampliam-se as formas de busca, mas em ambos o propósito de apresentar/divulgar produtos é o mesmo.

Esse cenário parece ser simples, porque se pode entender que o digital é independente de uma gênese que ocupa há tempos diferentes metodologias de teóricos, como

por exemplo, da Biblioteconomia. Também pode ser pouco, porque partindo destes princípios, caberia trazer uma grande variedade de linguagens que são entendidas por computadores. Porém, entre a compatibilidade de fazer o computador entender o que já é feito manualmente, há fundamentos que dão conta de práticas sedimentadas em teorias e/ou princípios.

Por isso, preferir a análise de um site, como o da *Amazon*, em um percurso que isenta representações com ontologias e respectivos formalismos faz parte da próxima seção. A relação com o pensamento biblioteconômico pela Catalogação de Assunto e a reflexão que ele suscita busca demonstrar uma configuração de formas que não são novas.

3 EM FOCO UMA ANÁLISE: O SITE DA AMAZON

Conforme indicado, para essa análise⁶ isenta-se o foco em ontologias e/ou formalismos (como por exemplo, *Resource Description Framework (RDF)*; *Simple Knowledge Organization System (SKOS)*; *Web Ontology Language (OWL)*; *Linked Data*), dentre outros). Da mesma forma, não há pretensão de vincular essa análise a Arquitetura da Informação. Delimitar o que não se pretende trabalhar é uma maneira de reconhecer a intencionalidade do que se vai fazer, mesmo considerando que no contexto da web semântica impera a necessidade de formalismos como ponte entre pessoas e computadores.

⁶ Análise realizada em junho de 2021 e, portanto, torna-se necessário levar em consideração o referido tempo de observação no site da Amazon.

Sendo assim, a análise consiste em olhar para o que o site da *Amazon* apresenta, especificamente em relação a representação, e demonstrar como os resultados aproximam-se dos fundamentos levantados na seção anterior. Essa aproximação pode trazer ou não convergências, porém mais importante é justificar fundamentos sedimentados em um espaço-temporal anterior ao digital.

Como forma ampla de resgate e, buscando um paralelo com a palavra formalismo, de acordo com Abbagnano (2007, p. 471), é “[...] toda doutrina que recorra à forma, em qualquer das significações do termo.” Esse mesmo autor, para o termo forma, discorre diferentes perspectivas chamando atenção para “[...] uma norma de procedimento [...]” (p. 470). Já norma, ainda segundo Abbagnano (2007, p. 716, *itálico do autor*), “[...] formou-se através da distinção e da contraposição entre o domínio empírico do *fato* (da necessidade natural) e o domínio racional do *dever ser* (da necessidade ideal). Sua validade não deriva do fato de ser ou não aceita ou aplicada, mas apenas do *dever ser* que exprime.”

Sendo assim, é possível unir a esse recorte o entendimento de que havendo ou não a implicação de aceitação, a existência de fundamentos para a representação com fins de recuperação é fato.

Voltando para os fundamentos apresentados na seção anterior, há um constructo pela via da representação. No site da *Amazon*, por exemplo, também há padrões e procedimentos de representação com o objetivo de dispor para as pessoas informação que identifique o ato de recuperar. Há nessa representação o objetivo de comunicação entre o

catálogo de produtos e pessoas interessadas em buscar informação, quer seja para fins de aquisição, comparação de preços, especificação de produtos, dentre outros motivos. Não há novidade nesse ponto.

Não se pode perder de vista que a organização de sites, com fins comerciais ou não, primeiro faz uso da organização por categorias, situação que antecede o ato de representação. Mas, pela segunda, ou seja, pela representação, chega-se nas formas de registro e disponibilização da informação. Na mesma medida de pensar as formas de representação é importante saber como a informação será buscada, pois uma exerce influência na outra. Isso há bastante tempo constitui tópico de interesse da Biblioteconomia.

Conforme demonstrado, a partir do catálogo dicionário três tipos de entradas (autor, título e assunto) são representados numa só lista alfabética. No site da *Amazon*, mesmo o resultado da busca fornecendo privilégio via patrocínio (ver figura 2), as formas de entrada dos produtos variam. Na figura 2 é possível observar que a busca realizada na categoria eletrodomésticos, com foco na palavra-chave refrigerador, recuperou primeiro a marca do produto *Hongzer*.

Figura 2: Resultado da busca no site da *Amazon* pela palavra-chave refrigerador

The screenshot shows the Amazon.com.br search results for 'refrigerador'. The search bar at the top contains the word 'refrigerador'. Below the search bar, there are navigation links for 'Todos', 'Venda na Amazon', 'Mais Vendidos', 'Prime', 'Livros', 'Lançamentos', 'Atendimento ao Cliente', and 'Moda da Amazon'. The search results are displayed in a grid format, showing four refrigerator models:

- Hongzer Mini geladeira USB**: A compact, white refrigerator with a USB port. Price: R\$ 5,999.00.
- Refrigerador French Door InverterQuattro 482L Preto**: A black French door refrigerator. Price: R\$ 87,40 de frete.
- Geladeira/Refrigerador Top Freezer 431L Platinum (TF555)**: A silver top freezer refrigerator. Price: R\$ 1,444.00.
- Geladeira/Refrige Defrost Electrolux Prático 240L Brn 127V**: A white compact refrigerator. Price: R\$ 1,444.00.

Fonte: site *Amazon*, pesquisa realizada em 20 de jun. 2021.

É possível confirmar o nome da marca *Hongzer*, que é a entrada do produto, conforme consta na figura 3. Por outro lado, a forma de entrada (figura 2) do segundo produto é refrigerador e dos seguintes geladeiras/refrigeradores. Nesse recorte constata-se ao menos duas formas de entrada para representação dos produtos no catálogo da *Amazon* e isso há bastante tempo foi difundido com *Cutter* no catálogo dicionário. Resta a dúvida da existência, ou não, de uma política que norteie a decisão das entradas nesse catálogo de produtos. Da mesma forma, há dúvida se existe padronização na forma de representação do(s) assunto(s). Na figura 3 é possível observar ‘informações sobre o produto’, também denominado no site de ‘detalhes técnicos’.

Figura 3: Seleção do primeiro produto que consta na figura 2:
Hongzer mini geladeira - 'Informações sobre o produto'

Informações sobre o produto

Detalhes técnicos

Nome da marca	Hongzer
Dimensões do produto	26 x 17.5 x 11.5 cm; 830 g
Eficiência	Perfect
Número da Peça	Hongzerx5mv0q1tyr
Configuração básica	Frigobares
Estantes	1
Peso	0.83 Quilogramas
Fabricante	Hongzer
Referência do fabricante	Hongzerx5mv0q1tyr
ASIN	B07Y33WK7M

Fonte: site *Amazon*, pesquisa realizada em 20 de jun. 2021.

Na figura 4 é possível notar a 'descrição do produto', e especificamente em 'descrição', nota-se indicação de um breve resumo. Já 'especificações' contribuem com as 'informações sobre o produto' (figura 3). As representações dos campos das figuras 3 e 4 não são recuperáveis, mas nem por isso deixam de ser informativas.

Figura 4: Seleção do primeiro produto que consta na figura 2:
Hongzer mini geladeira - 'Descrição do produto'

Descrição do produto

Descrição:
Esta mini geladeira tem aparência elegante e legal, adequada para colocar bebidas enlatadas, cerveja, leite, bebidas, etc. Alimentado por cabo USB, adequado para todos os computadores, carros ou outras interfaces USB, sem necessidade de pilhas. Para que você possa desfrutar de bebidas frias no escritório ou no carro a qualquer momento, muito conveniente e prático. Ele também pode servir como aquecedor mantendo as bebidas quentes por mais tempo.

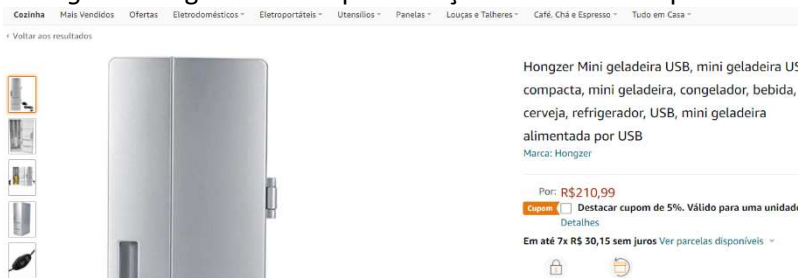
Especificações:
Condição: 100% novo
Material: ABS
Cor: como mostrado na imagem.
Tamanho externo: aprox. 8,5 x 12 x 25 cm.
Comprimento do cabo USB: aprox. 1,2 m.
Comprimento do cabo USB operacional: aprox. 1,0 m.
Consumo de energia da refrigeração: 5 V 1000 mA
Consumo de energia de aquecimento: 5 V 500 mA
Temperatura de aquecimento mais alta: 50 °C.
Temperatura de resfriamento mais baixa: 10 °C
Tempo de refrigeração: 10 ~ 15 minutos.
Tempo de aquecimento: longa duração.
Ruído: 35DB
Sistema operacional: Compatível com IBM PC/Mac
Peso do pacote: 820 g
Quantidade: 1 peça

Incluído na embalagem:
1 mini geladeira USB.

Fonte: site *Amazon*, pesquisa realizada em 20 de jun. 2021.

O que consta nas figuras 3 (detalhes técnicos) e 4 (especificações, pois descrição equivale a um breve resumo), de forma ampla, alcançam elementos descritivos do produto. A representação do assunto consta na figura 2, logo abaixo da imagem do produto. Para melhor visualização, segue a seleção do produto na figura 5.

Figura 5: Seleção do primeiro produto que consta na figura 2:
Hongzer mini geladeira - representação de assunto do produto



Fonte: site *Amazon*, pesquisa realizada em 20 de jun. 2021.

É possível inferir que: Hongzer Mini geladeira USB, / mini geladeira USB compacta, / mini geladeira, / congelador, / bebida, / cerveja, / refrigerador, / USB, / mini geladeira alimentada por USB – são os assuntos que representam o produto e aproximam-se dos cabeçalhos de assunto. Nas regras de Cutter, os cabeçalhos têm relação direta com a entrada no catálogo. Ao pesquisar 'mini geladeira USB compacta', nota-se a recuperação de vários produtos com esse cabeçalho, conforme é possível observar na figura 6.

Não há presença de listas de cabeçalhos de assunto, mas há termos, palavras, frases, nomes ou expressões empregadas. Conforme já apresentado, Cutter (1904) também definiu cabeçalho (*heading*) como a palavra por ordem alfabética que fornece o lugar da entrada no catálogo e pode

ser determinado pelo nome do autor, do assunto ou palavra do título.

Figura 6: Pesquisa com o cabeçalho ‘mini geladeira USB compacta’



Fonte: site *Amazon*, pesquisa realizada em 20 de jun. 2021.

Ainda é possível considerar que os Objetivos do catálogo, conforme estabelecido por Cutter (1904), convergem no contexto do site da *Amazon*, pois: permite a uma pessoa encontrar um produto no qual é conhecida a marca, um título (nome do produto) e o assunto; mostra o que o site possui por uma marca determinada, com um assunto determinado e o tipo de produto; e ajuda na escolha de um produto quanto a sua edição (ano) e o seu caráter (considerando para isso os detalhes técnicos e a descrição do produto).

Dessa forma, no primeiro objetivo, pela identificação de um produto no catálogo é possível alcançar o segundo objetivo, que mostra os produtos que o site tem, possibilitando o agrupamento de atributos dos produtos. Já o terceiro objetivo ajuda a escolher um produto por sua edição, ou particularidades, pois quem busca encontrar produtos

similares necessitará realizar uma seleção, por exemplo, entre as várias edições disponíveis. No cenário das *Rules*, no século XIX, não havia abundância de muitas edições dos livros, mas foi ponto merecedor de atenção. Atualmente, para a farta quantidade de produtos no catálogo de um site de comércio eletrônico, como no da *Amazon*, torna-se atual e relevante recuperar esses Objetivos.

Outra convergência observada é a presença do uso de remissivas, conforme pode ser visualizado na figura 7.

Figura 7: Apresentação da remissiva ‘veja também’



Fonte: site *Amazon*, pesquisa realizada em 20 de jun. 2021.

Ao selecionar a categoria cozinha e realizar a busca (clcando na imagem da lupa), o resultado é um índice em ordem alfabética e a presença de uma remissiva na forma ‘veja também’. Os seis meios, segundo Fiúza (1987, p. 49), para alcançar os Objetivos, que ela denominou de “Objetos”, e em interseção na figura 1 (seção anterior- referencial teórico), estão presentes no site da *Amazon*: “1. Entradas de autor, com as referências necessárias; 2. Entradas de título ou referências ao título; 3. Entradas de assunto, referências cruzadas; 4. Entrada de forma e língua; 5. Transcrição da edição e impronta, com notas, quando necessário; 6. Notas especiais.” Entende-se entrada de autor como entrada pela marca; entrada de

forma a relação com os detalhes técnicos e a descrição do produto, mas apesar desses campos não serem recuperáveis na forma de *links*, são informativos; transcrição da edição o ano do produto.

Em alguma medida, os Objetivos aproximam-se dos Meios, e os aspectos principais são estabelecidos pelos atributos da marca, do título do produto e assunto que moldam a forma da entrada do catálogo de produtos. Mas, conforme já apresentado, a preocupação central em Cutter não foi com o assunto e sim com a sua forma de entrada.

Formas para a 'Entrada' no catálogo já haviam sido previstas por Panizzi, mas o 'Estilo' e os Objetivos tomam significância por propor critérios pragmáticos para uma política do catálogo. Firma-se proximidade à Entrada e a busca no catálogo, quer seja o catálogo de biblioteca ou o catálogo de produtos, como no site da *Amazon*.

4 APONTAMENTOS FINAIS

Ilustrou-se, de maneira pontual, uma análise permitindo a visualização do cenário digital, mas que não exige a responsabilidade de buscar sustentação teórica no passado, no que ele é. O olhar não pode ser apenas prático para a solução de problemas.

A consolidação de fundamentos e princípios para qualquer área é resultado de convergências e divergências e, assim, entender o lugar de origem e/ou possíveis influências teóricas é importante para não haver multiplicação de estruturas sem conhecimento da origem.

O cenário web demonstra desafios aos cursos de Biblioteconomia e, por vezes, é tratado com prioridade no ensino. Docentes e/ou discentes podem olhar com satisfação esse cenário, mas frente aos efeitos do recorte digital há fundamentos teóricos que se tornam atuais, conforme demonstrado na análise do site de comércio eletrônico. Ao invés de priorizar um contexto em detrimento do outro, o ganho está em conjugá-los. Ademais, o entendimento da teoria quando antecedido à prática sustenta o fazer e a atuação profissional em contextos diversificados.

Articular os fundamentos teóricos ao ensino, à pesquisa e ao mercado de trabalho com temáticas relativas à organização da informação, quer seja pela perspectiva da Catalogação de Assunto e/ou outras, contribui para o fortalecimento da Biblioteconomia e da Ciência da Informação.

REFERÊNCIAS

ABBAGNANO, N. *Dicionário de filosofia*. 5. ed. rev. e ampl. São Paulo: Martins Fontes, 2007.

BUONOCORE, D. *Diccionario de Bibliotecologia*. 2. ed. aum. Buenos Aires: Marymar, 1976.

CUTTER, C. A. *Rules for a printed dictionary catalogue*. Washington: Librarian of the Boston Athenaeum, 1876.

Disponível em:

<https://archive.org/details/cu31924029518978>. Acesso em: 2 nov. 2016.

CUTTER, C. A. *Rules for a printed dictionary catalogue*. 4. ed. Washington: Government Printing Office, 1904.

FARANI, C. Durante pandemia Amazon tem o maior trimestre de vendas online de todos os tempos. *Forbes*, São Paulo, 10 ago. 2020. Disponível em: <https://forbes.com.br/forbes-collab/2020/08/camila-farani-durante-pandemia-amazon-tem-o-maior-trimestre-de-vendas-online-de-todos-os-tempos/>. Acesso em: 20 jun. 2021.

FIÚZA, M. M. A catalogação bibliográfica até o advento das novas tecnologias. *Revista da Escola de Biblioteconomia da UFMG*, Belo Horizonte, v. 16, n. 1, p. 43-53, mar. 1987.

GARRIDO ARILLA, M. R. *Teoría e historia de la catalogación de documentos*. Madrid: Síntesis, 1999.

MANN, M. *Catalogação e classificação de livros*. Rio de Janeiro: Editôra Fundo de Cultura, 1962.

MIKSA, F. L. The legacy of the library catalogue for the present. *Library Trends*, Champaign, v. 61, n. 1, p. 7-34. 2012.

OLSON, H. A. The power to name: representation in library catalogs. *Journal of Women in Culture and Society*, Chicago, v. 26, n. 3, p. 639-668. 2001.

SHERA, J. H.; EGAN, M. E. *Catálogo sistemático: princípios básicos e utilização*. Brasília: Editôra Universidade de Brasília, 1969.

STROUT, R. F. The development of the catalog and cataloging codes. *The Library Quarterly*, Chicago, v. 26, n. 4, p. 254-275, oct. 1956.

TAYLOR, M. S. *Fundamentals of practical cataloguing*. London: George Allen and Unwin, 1948.

WORLD ENCYCLOPEDIA OF LIBRARY AND INFORMATION SERVICES. 3. ed. Chicago: American Library Association, 1993.

CAPÍTULO 8

A INSTÂNCIA DE ASSUNTO NO CONTEXTO DIGITAL: EM CENA O TRATAMENTO TEMÁTICO DA INFORMAÇÃO

Lais Pereira de Oliveira

1 INTRODUÇÃO

Assunto, de acordo com o Dicionário de Biblioteconomia e Arquivologia, é a matéria tratada em um documento (Cunha; Cavalcanti, 2008). Diz respeito, portanto, a um conteúdo sobre o qual se discorre. Barité *et al.* (2015) definem-no como tópico ou tema documental. Embora configurado como um conceito escorregadio (Lancaster; Elliker; Connell, 1989) deixado, de certo modo, em segundo plano (Guedes; Martinho; Moraes, 2009), na Ciência da Informação (CI), a identificação e descrição do conteúdo documental tem lugar na classificação, na indexação e na catalogação de assunto (Naves, 2001).

Assim, assunto representa uma preocupação primeva em Biblioteconomia e Ciência da Informação. Grosso modo, o olhar sobre a instância temática de um documento se estabeleceu juntamente com a clássica dedicação bibliotecária de descrever atributos formais de produção de uma dada obra.

Todavia, identificar assunto é algo problemático (Hutchins, 1977), de difícil parametrização e, que representa uma lacuna na representação da informação (Guedes;

Martinho; Moraes, 2009), posto que, enquanto ação intelectual (Lima; Maculan, 2014; Redigolo, 2014), conjuga subjetividade, complexidade e um caráter desafiador (Dal'Evedove; Fujita, 2013; Dias; Naves; Moura, 2001; Fujita, 2003; Lima; Maculan, 2014; Sousa, 2013). Desse modo, a atividade de tratamento temático da informação (TTI) (Braz; Carvalho, 2017; Sousa; Fujita, 2013) pressupõe esforços pontuais em prol dessa operação de trato do conteúdo documental.

Ao viabilizar a determinação do conteúdo, em sua preocupação de possibilitar acesso ao assunto contido em um documento (Sousa, 2013), o TTI reveste-se de um caráter seminal. Onde se pode inferir seu aspecto nuclear (Guimarães, 2008), na evidência do teor documental, que não se restringe, porém, às bibliotecas e seus acervos físicos, estendendo-se às bibliotecas digitais (Dias, 2004) e a outras estruturas, plataformas e ambiências, incluindo-se aqui os repositórios digitais, os portais institucionais e as mídias sociais, em sua configuração temática via rótulos de assunto.

A dimensão temática consubstancia-se, afinal, na identificação do conteúdo de documentos, assim como de bancos de dados (Lancaster; Elliker; Connell, 1989). Entendendo, pois, que a dinamicidade de trato por assunto estende-se ao contexto digital, há que se examinar pontualmente esse entorno. Mesmo ao se considerar a integração entre descrição física e organização do conteúdo, tornada viável pelo desenvolvimento tecnológico (Ortega, 2008).

Nesse sentido, objetiva-se discorrer teoricamente sobre o tratamento temático da informação no contexto digital. Ato contínuo parte-se da problemática delimitação da instância de assunto em Biblioteconomia e Ciência da Informação, tanto quanto de seu caráter de tratamento, ainda com lacunas teóricas e metodológicas (Sousa, 2013) a serem mais bem delimitadas e explicitadas, demandando enriquecimento de seu quadro teórico (Sales, 2012). Considera-se, ainda, a complexidade em organização da informação, no contexto da Internet (Pando, 2018).

Nesse intento, desenvolveu-se pesquisa de natureza teórica e abordagem qualitativa, prioritariamente estabelecida sobre artigos de periódicos e capítulos de livro versados no tema tratamento temático da informação. Tais materiais foram obtidos a partir de busca por assunto com a expressão supracitada na Base de Dados Referencial de Artigos de Periódicos em Ciência da Informação (BRAPCI), bem como pela seleção de livros seminais a respeito de tal dimensão temática. Ato contínuo, esse cerne de discussão pode agregar aos estudos produzidos em TTI e ao hiato teórico-conceitual sobre a instância de assunto, em sua iminência no contexto digital.

2 A INSTÂNCIA DE ASSUNTO EM BIBLIOTECONOMIA E CIÊNCIA DA INFORMAÇÃO

A instância de assunto, configurada como a informação relevante de um texto (Fujita, 2003) e, a conseqüente operação de análise do conteúdo documental, foram, historicamente, negligenciadas (Hutchins, 1977). Por simples observação se constata o quão importante é e o quanto se faz

presente o assunto, tema ou tópicos, seja em tarefas cotidianas, de trabalho e estudo, ou mesmo na crucialidade em representar o teor de um documento para acesso futuro em bases de dados e sistemas de informação.

Considera-se, aqui, dentro da conjuntura profissional supracitada, a dificuldade em seu desenvolvimento aplicado, tanto quanto em sua prática de ensino (Naves, 2001). Ora, mais do que parâmetros especificamente definidos para que o bibliotecário possa representar tematicamente um documento, há que se atentar até que ponto se consegue lecionar sobre a instância de assunto e as diretrizes aplicadas em sua sistematização, nas graduações em Biblioteconomia.

Ademais, em que pese ser a organização da informação um dos pontos de preocupação central na Ciência da Informação, inclusive justificando a emergência do campo, acaba-se por ampliar o espectro de atenção para com a instância de assunto, aqui colocado. Isto em condições de incutir ideias que permitam enxergá-la de modo expressivo e amplo juntamente com outras temáticas correlatas, como a catalogação descritiva, da mesma forma que em sua condição limítrofe com tantos outros processos e métodos estudados e problematizados na CI.

Fundamentalmente, além da pouca atenção recebida por essa atividade analítica de trato temático (Lima; Maculan, 2014), de caráter intrínseco (Lancaster; Elliker; Connell, 1989; Guimarães, 2003), o próprio termo “assunto” tem sido utilizado de modo bastante vago (Hickey, 1976), na área de Biblioteconomia e Ciência da Informação. A alternância e a diversidade na sua nomeação se juntam, pois, à escassa

abordagem de forma individualizada e explícita, que mutuamente influi na diminuta menção a esse eixo tão fundamental que é o do teor documental.

Em bases gerais, a preocupação com a instância de assunto foi enfatizada a partir dos estudos de Cutter (Guedes; Martinho; Moraes, 2009). Teóricos clássicos como Bliss, Kaiser, Shera e Foskett também contribuíram grandemente nessa especificação, assim como os trabalhos do *Classification Research Group* (Hider, 2020; Oliveira; Grácio; Martínez-Ávila, 2020; Sales, 2012). Todavia, tais incursões precisam ser ainda maiores, como forma de balizar a compreensão a respeito desse elemento e explorar metodologias para o trato analítico documental por assunto, considerando também o contexto digital.

No que se refere ao trato informacional via conteúdo, é grandemente expresso na literatura como análise de assunto (Fujita, 2003). Mas, como constatado em estudo anterior (Oliveira; Grácio; Martínez-Ávila, 2020), tratamento temático da informação e, representação temática são, do mesmo modo, expressões frequentemente empregadas para designá-lo dentro do contexto brasileiro. Imprime-se, portanto, uma vasta conjuntura designativa.

Destarte, o tratamento temático da informação remonta a práticas desenvolvidas na Antiguidade. Via de regra, mediante construção de listas e registros da autoria dos documentos e, ainda, dos conteúdos e de outras informações complementares presentes nas tabuletas de argila (Fujita, 2003; Medeiros, 2019; Simões, 2017). Além do caráter aplicado do TTI, a própria instância de assunto é objeto de

análise na Biblioteconomia (Hjørland, 2017) e, seu tratamento, extremamente relevante para a prática bibliotecária, tanto quanto na conceituação em Ciência da Informação. Apesar disso, o TTI é pouco abordado nas pesquisas (Redigolo, 2014). Além do que, os estudos da área costumam se concentrar na etapa de tradução (Dias, 2004), deixando descoberta a etapa analítica do conteúdo.

Fujita (2003) explica que, na Antiguidade, em determinado momento foi demandada uma ordenação documental por assunto. Ato contínuo, essa necessidade se estendeu ao longo do tempo, atrelada ao desenvolvimento científico e ao surgimento de novas áreas do conhecimento.

A dinamização dos assuntos tornou-se, pois, uma frente de preocupação bibliotecária, no que se refere à forma de descrevê-los e melhor representá-los. Embora não tenha se exaurido e, persista como um desafio à profissão. Notadamente, o trato de assunto é, agora, desafiado frente ao contexto digital, tanto quanto foi em relação aos acervos e registros do conhecimento em suporte impresso tradicional.

2.1 Tratamento temático da informação

Descrição física e temática, correspondentes à forma e ao conteúdo, proporcionam a estruturação da representação documentária (Sousa, 2013). Pode-se explicitar que essa perspectiva dual de descrição, representação ou trato informacional (Catarino; Cervantes; Andrade, 2015; Sousa, 2013), haja vista que todas essas variantes designativas estão presentes, propicia a busca e seleção da informação.

Desse modo, o conjunto de representações descritivas e temáticas reflete o caminho para ascender às informações. Em outras palavras, o trabalho bibliotecário em prol do tratamento descritivo e temático visa recuperar o registro de informação, o qual remete ao documento e possibilita seu acesso (Sousa, 2022).

A especificidade do tratamento temático da informação está no fato de se dedicar ao assunto, estabelecendo formas de evidência deste, tanto sob forma notacional quanto verbal. Nesse ensejo, também incita a construção de instrumentos lingüístico-terminológicos e normativo-descritivos, os quais amparam a ação de descrição do teor documental.

Para além da esfera de seus processos e produtos, o alcance do tratamento temático da informação incide sob outras vias. Lancaster, Elliker e Connell (1989, p. 35, tradução nossa) explicitam que, “do ponto de vista do usuário, a análise de assunto está intimamente ligada ao acesso do assunto”. A organização dos materiais informacionais busca, de fato, alcançar a informação neles contemplada (Sousa; Fujita, 2013), especialmente pelo teor que congregam, razão pela qual o contributo do TTI se estende para além do nível da representação, alcançando a esfera da busca e recuperação da informação.

Assim, no caso do tratamento temático da informação alcança-se a determinação dos termos de indexação e a representação via notação (Braz; Carvalho, 2017), abarcando, portanto, o conteúdo temático (Monteiro; Giraldes, 2008) contido na informação. O teor documental é expresso, ainda,

pela condensação textual, que toma forma na elaboração de resumos.

Nesse ensejo, o TTI atua como mediador entre documento e usuário (Braz; Carvalho, 2017), tornando-se uma necessidade (Redigolo, 2014) que ganha força conforme o conteúdo documental se torna passível de socialização (Sousa, 2013). Além do que, em sua essência, o TTI é a ponte para acesso à informação por assunto, retroalimentando a prática de pesquisa, busca e seleção das informações.

Há, entretanto, uma forte dimensão teórico-conceitual que ganha corpo no contexto científico brasileiro, principalmente por conta de o tratamento temático da informação ser, no Brasil, como observam Gracioso, Martínez-Ávila e Simões (2019), um verdadeiro objeto de problematização. Por essa razão o TTI se torna cerne de teorizações e conceituações, mais do que dimensão aplicada de representação do assunto, que tem lugar pelas mãos bibliotecárias.

Destacadamente, o TTI foi impulsionado nos anos de 1970, com a publicação de *“The subject approach to information”*, traduzida no Brasil como *“A abordagem temática da informação”*, tendo refletido, ainda, no ensino de Biblioteconomia brasileiro (Guimarães, 2008). Atinge, ademais, os propósitos de organização dos acervos e de acesso à informação por assunto (Braz; Carvalho, 2017), como supramencionado. Todavia, o TTI é mais do que processos (Oliveira; Martínez-Ávila, 2019). Em verdade, ocupa relevante espaço na literatura e na prática profissional (Guimarães, 2008).

Em âmbito aplicado, é preciso ressaltar que há uma confusão terminológica nos processos englobados no TTI (Sousa; Fujita, 2013), assim como em sua própria forma designativa conforme já mencionado, em razão das diversas correntes que lhe constituíram e resultaram em distintas concepções em seu entorno (Guimarães, 2008). Não por acaso, “a complexidade existente na identificação e representação do tema de um documento também causa a multiplicação de nomenclaturas e definições presentes na literatura” (Guedes; Martinho; Moraes, 2009, não paginado).

Em que pese sua inconsistência designativa, o tratamento temático da informação viabiliza a descrição do assunto abordado no documento (Sousa, 2013), considerando as distintas finalidades e meios de recuperação da informação, nas mais diversas fontes (Lancaster; Elliker; Connell, 1989; Sousa; Fujita, 2013), assim como uma variedade de elementos que permeiam a representação do conteúdo. Mormente:

No que se refere ao conteúdo a representação temática, auxilia significativamente na recuperação da informação a partir da determinação e disponibilização de pontos de acesso. (Catarino; Cervantes; Andrade, 2015, p. 106)

Não obstante, por um longo período a análise temática foi tida como qualidade de alguns profissionais (Sousa, 2013). O contexto digital, contudo, veio a alterar ainda mais esse cenário, na imprescindibilidade de novos formatos de descrição (Castro; Santos, 2009), estendendo-se também às ações do usuário da informação, produtor de conteúdos nesse meio, demandando adequações nos instrumentos de

representação e organização a partir de novas estruturas e ferramentas tecnológicas (Arakaki; Santarém Segundo; Vidotti, 2020). Incitam-se, pois, modelos de atribuição colaborativa de rótulos de assunto, discute-se a pertinência de assuntos tradicionalmente contemplados em descritores com controle linguístico e, estruturados em instrumentos terminológicos e classificatórios, além de se questionar, cada vez mais, a padronização na descrição do conteúdo, classicamente adotada na área de Biblioteconomia.

2.2 O tratamento temático da informação no contexto digital

A manipulação de informação já estabelecida em âmbito digital é crescente. Como destacam Arakaki, Santarém Segundo e Vidotti (2020) acessar recursos digitais representa tarefa instantânea, amparada em variadas formas de tecnologia e em conceitos da organização da informação.

Não obstante, o expressivo aumento da produção de conteúdos e, inclusive, da informação científica, consubstanciada em áreas interdisciplinares, representa desafio substancial à atividade de organização e tratamento das informações, ao qual se soma a própria complexidade da análise de assunto (Naves, 2001). Abreu e Monteiro (2010, p. 9) chegam a destacar que “a automação do tratamento da informação apresenta diferentes paradigmas na organização dos estoques informacionais”.

De forma específica, o TTI é grandemente impactado frente à multiplicidade e diversidade de assuntos existentes, à especialização dos campos do conhecimento e à interlocução

relacional e interdisciplinar entre muitos deles. Para além disso, novas estruturas e plataformas desafiam a articulação das representações documentais mediante seu teor, em se tratando das formas de expressão de assuntos que antes se estabeleciam embasadas em concepções clássicas sobre os diversos temas e, da tendência bibliotecária de padronização e universalização dos arranjos de conteúdos, que passam a ser rediscutidas e repensadas.

Há que se ter em mente, também, que o assunto é identificado a partir de um enunciado temático (Fujita, 2003). Assim, o TTI precisa contribuir para a adequada identificação do conteúdo (Sousa, 2013), exprimindo adequadamente o teor documental, por mais complexa que essa atividade possa parecer. Decorre, porém, a ausência de parametrização e método na análise de assuntos já presente na conjuntura analítica estabelecida sobre o documento impresso e que ora se estende ao conjunto nato digital.

Todavia, a dinamicidade em TTI torna-se ainda maior ao se considerar o contexto digital. Afinal, mais informações são publicadas, armazenadas e disponibilizadas na *web* (Castro; Santos, 2009), configurando uma verdadeira explosão de conteúdos *online* (Guimarães; Quaresma, 2020). Ademais, as tecnologias da *web* impactaram na forma de representar informação, com efeito nos sistemas de recuperação da informação (Lima, 2018).

De modo que, além de tal processo representar um dos maiores desafios no entorno da organização da informação, pela dinamicidade envolvida na análise e descrição de assunto, há que se considerar a quantidade de materiais com os quais é

preciso lidar hodiernamente (Guedes; Martinho; Moraes, 2009), potencializada nos documentos produzidos em meio digital, na forma de artigos, e-books, *pré-prints* e etc, assim como os conteúdos gerados pelo próprio usuário, fora do contexto acadêmico-científico.

Destarte, embora seja apoiado pelos recursos tecnológicos, o tratamento temático da informação (Catarino; Cervantes; Andrade, 2015) é grandemente impactado pelo contexto digital, na robusta estrutura informacional que nele têm lugar, cujo parâmetro organizativo não cresceu na mesma proporção. Sobretudo porque a crescente disponibilização de conteúdos digitais convive com a ausência de descrição padronizada de tais recursos (Castro; Santos, 2009). Ademais, em que pese o avanço tecnológico, o fator humano é fundamental na abstração e distinção de um assunto (Santos, 2017), deixando sempre uma lacuna no que se refere à automatização dos processos de representação nesse âmbito.

Fato é que a análise de assunto interfere diretamente na recuperação da informação (Redigolo, 2014). E, em que pese o desenvolvimento de estruturas taxonômicas e ontológicas, assim como de todo um aparato pautado na web semântica, uma lacuna substancial persiste no entorno representativo em meio digital, especialmente na evidência temática dos distintos conteúdos gerados.

A web semântica proporciona melhorias no trato informacional do conteúdo presente no contexto digital (Arakaki; Santarém Segundo; Vidotti, 2020). Iniciativas dessa ordem possibilitam avançar sobre representações mais

consistentes considerando a semântica e o significado, embora o problema não se restrinja a isso.

Mormente, há que se considerar a existência de inúmeros documentos armazenados em sistemas *online*, cuja recuperação se faz igualmente necessária (Braz; Carvalho, 2017). Todavia, uma pequena parcela deles conta com uma especificidade no trato de assunto. Geralmente aqueles armazenados em ambientes específicos (Castro; Santos, 2009), fazendo com que uma série de outros sequer tenham existência conhecida. De modo que:

essa condição facilitada e simples de recuperá-los não se efetiva na prática, uma vez que o meio digital demanda, também, acurados processos de representação e arranjo para adequada visualização, pesquisa e seleção das informações que congrega. (Oliveira; Souza, 2021, p. 12)

Como destacam Arakaki, Santarém Segundo e Vidotti (2020), nos sistemas de informação, o acesso a recursos de informação digitais é quase que automático. Afinal, metadados descritivos e temáticos cuidadosamente estruturados viabilizam a evidência do atributo informacional. Isso em se tratando de conjuntos documentais formalmente armazenados em sistemas de recuperação informação.

Ora, várias coleções nascem digitais (Weitzel; Santos, 2020), exigindo formas de tratamento já em seu fluxo de produção. Outro aspecto marcante diz respeito à facilidade de propagação dos acervos de bibliotecas, pela via das informações documentárias, disponibilizadas no catálogo *online* (Sousa, 2013), permitindo a formação de redes (Lima, 2018). E, embora catálogos de bibliotecas apresentem

limitações na pesquisa por assunto (Lancaster; Elliker; Connell, 1989), cumprem importante papel na disseminação de metadados, inclusive temáticos, mas, que também demandam aperfeiçoamento.

Entretanto, os catálogos não indexam toda a informação produzida. Ato contínuo, uma diversidade de conteúdos imagéticos, sonoros e textuais ganha corpo no uso cotidiano de portais, mídias e outros canais digitais. Materiais de toda ordem são produzidos e difundidos por esse meio, sem que se estabeleçam adequados processos de descrição. As bolhas criadas em estruturas de mídias sociais apregoam tagueamento de conteúdo enviesado para alavancar os acessos e visualizações, em detrimento da confiabilidade e solidez da fonte e da categorização para ela estabelecida. Desinformações se estabelecem e ampliam seu alcance na Internet, por mais absurdo que seja seu teor.

De forma ampla, conglomerados de dados se juntam a informações pré-estruturadas na *web*, um conjunto expressivo destas, porém, ofuscada por interesses comerciais na evidência do conteúdo que congregam. Um entorno armazenado em repositórios, bases de dados e catálogos de bibliotecas também tem lugar e, ainda que imerso em problemas de recuperação por assunto, encontra uma conjuntura mais favorável a esse objetivo.

Redigolo (2014) lembra dos esforços normalizadores e de sua importância para a análise de assunto, em função da subjetividade e inferência que cercam essa atividade de trato do conteúdo. Contudo, trata-se aqui da perspectiva bibliotecária de atuação no tratamento temático da

informação, à qual se deve incluir a do usuário, que participa cada vez mais da produção de distintos conteúdos na *web* e de sua incursão, nesse meio, sem regulação e padronização.

Mormente, enquanto conjunto de ações vinculadas à organização do conhecimento e da informação (Gracioso; Martínez-Ávila; Simões, 2019), o TTI acaba ficando restrito ao contexto documentário, deixando passar uma enormidade de questões. Inclusive a instituição de uma representação de assunto cada vez mais desenvolvida pelas mãos do produtor de conteúdo na *web* (Oliveira; Souza, 2021), sobre a qual a dimensão temática precisa, também, se debruçar, em busca de novas metodologias e soluções.

3 CONCLUSÃO

Ao buscar discorrer teoricamente sobre o tratamento temático da informação no contexto digital, exploraram-se aspectos relacionados à amplitude de informações nesse meio e à diversidade de conteúdos produzidos pelo usuário e, portanto, não exclusivo aos acervos em repositórios, bases de dados e catálogos de bibliotecas. Além disso, destacou-se a descontinuidade do parâmetro organizativo em meio digital, que não acompanhou a crescente estrutura informacional nesse âmbito, com padrões de descrição suficientes.

Especificaram-se, ainda, a dualidade entre fator humano e automatização nos processos de representação da informação em meio digital, com a conseqüente lacuna na evidência temática das informações, bem como a multiplicidade de formatos e ambiências cuja estruturação não

reflete a confluência existente em sistemas de informação estruturados.

Destacadamente, evidenciaram-se, também, a demanda organizativa em acompanhamento ao fluxo de produção do documento digital e, a necessidade de aperfeiçoamento no trato informacional, mesmo nos catálogos e repositórios de bibliotecas. E, por fim, destacaram-se a diversidade de formatos em meio digital, a dinamicidade das bolhas digitais e da desinformação, os interesses comerciais por trás da evidência do conteúdo nesse âmbito, a mescla de conglomerados de dados e de informações pré-estruturadas na *web* e, a necessidade de inclusão da perspectiva organizativa do usuário.

As questões apresentadas permitem refletir sobre o desafio permanente acerca da instância de assunto, na área de Biblioteconomia e Ciência da Informação. Ato contínuo, envolvendo também seu prospecto de tratamento, seminal e fundante, cada vez mais propício a novas problematizações.

REFERÊNCIAS

- ABREU, J. G de; MONTEIRO, S. D. Matrizes da linguagem e a organização virtual do conhecimento. *Ciência da Informação*, Brasília, v. 39, n. 2, p. 9-26, maio/ago. 2010.
- ARAKAKI, A. C. S.; SANTARÉM SEGUNDO, J. E.; VIDOTTI, S. A. B. G. Aplicações tecnológicas em representação e organização da informação. *In: SIMÕES, Maria da Graça; LIMA, Gercina Ângela de (coord.). Do tratamento à organização da informação: reflexões sobre concepções, perspectivas e*

tendências. Coimbra: Imprensa da Universidade de Coimbra, 2020. p. 243-266.

BARITÉ, M.; COLOMBO, S.; DUARTE BLANCO, A.; SIMÓN, L.; CABRERA CASTROMÁN, G.; ODELLA, M. L.; VERGARA, M. *Diccionario de organización del conocimiento: clasificación, indización, terminología*. 6. ed. corr. aum. Montevideo: CSIC, 2015.

BRAZ, M. I.; CARVALHO, E. S. de. Práticas em tratamento temático da informação: interfaces de ensino e aprendizagem. *Revista Brasileira de Biblioteconomia e Documentação*, São Paulo, v. 13, n. esp., 2017.

CASTRO, F. F. de; SANTOS, P. L. V. A. da C. Uso das tecnologias na representação descritiva: o padrão de descrição bibliográfica semântica MarcOnt Initiative nos ambientes informacionais digitais. *Ciência da Informação*, Brasília, v. 38, n. 1, p. 74-85, jan./abr. 2009.

CATARINO, M. E.; CERVANTES, B. M. N.; ANDRADE, I. A. de. A representação temática no contexto da web semântica. *Informação & Sociedade: Estudos*, João Pessoa, v. 25, n. 3, p. 105-116, set./dez. 2015.

CUNHA, M. B. da; CAVALCANTI, C. R. de O. *Dicionário de Biblioteconomia e Arquivologia*. Brasília: Briquet de Lemos / Livros, 2008.

DAL'EVEDOVE, P. R.; FUJITA, M. S. L. Estudo sociocultural da comunidade discursiva do tratamento temático da informação em bibliotecas universitárias. *Encontros Bibli: Revista Eletrônica de Biblioteconomia e Ciência da Informação*, Florianópolis, v. 18, n. 36, p. 23-50, 2013.

DIAS, E. W. Análise de assunto: percepção do usuário quanto ao conteúdo de documentos. *Perspectivas em Ciência da*

Informação, Belo Horizonte, v. 9, n. 2, p. 146-157, jul./dez. 2004.

DIAS, E. W.; NAVES, M. M. L.; MOURA, M. A. O usuário-pesquisador e a análise de assunto. *Perspectivas em Ciência da Informação*, Belo Horizonte, v. 6, n. 2, p. 205-221, jul./dez. 2001.

FUJITA, M. S. L. A identificação de conceitos no processo de análise de assunto para indexação. *Revista Digital de Biblioteconomia e Ciência da Informação*, Campinas, v. 1, n. 1, p. 60-90, jul./dez. 2003.

GRACIOSO, L. S.; MARTÍNEZ-ÁVILA, D.; SIMÕES, M. da G. de M. “Tratamento Temático da Informação” na pesquisa Brasileira em Ciência da informação: percursos e relações. *Scire*, Zaragoza, v. 25, n. 2, p. 23-34, jul./dic. 2019.

GUEDES, E. G. F.; MARTINHO, N. O.; MORAES, J. B. E. de O assunto na Ciência da Informação: a questão do aboutness. *In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO*, 10., 2009. *Anais [...]*. João Pessoa: UFPB, 2009.

GUIMARÃES, F. X.; QUARESMA, J. C. da S. Representação e disseminação dos conteúdos digitais: relato de experiência da catalogação das *lives* do Projeto Registros de Informação (SIBI/UFBA). *In: SEMINÁRIO DE PESQUISAS FESPSP*, 9., 2020, São Paulo. *Anais [...]*. São Paulo: FESPSP, 2020.

GUIMARÃES, J. A. C. A análise documentária no âmbito do tratamento da informação: elementos históricos e conceituais. *In: RODRIGUES, Georgete Medleg; LOPES, Ilza Leite (org.). Organização e representação do conhecimento na perspectiva da Ciência da Informação*. Brasília: Thesaurus, 2003. p. 100-117. (Estudos Avançados em Ciência da Informação, v. 2).

GUIMARÃES, J. A. C. A dimensão teórica do tratamento temático da informação e suas interlocuções com o universo científico da International Society for Knowledge Organization (ISKO). *Revista Ibero-Americana de Ciência da Informação*, Brasília, v. 1, n. 1, p. 77-99, jan./jun. 2008.

HICKEY, D. J. Subject analysis: an interpretive survey. *Library Trends*, Champaign, v. 25, p. 273-291, jul. 1976.

HIDER, P. Origins of the knowledge organization field. In: HJØRLAND, B.; GNOLI, C. (ed). *ISKO Encyclopedia of Knowledge Organization*. Baden-Baden: ISKO, 2020.

HJØRLAND, B. Subject (of documents). *Knowledge Organization*, Würzburg, v. 44, n. 1, p. 55-64, 2017.

HUTCHINS, W. J. On the problem of 'aboutness' in document analysis. *Journal of Informatics*, Norwich, v. 1, n. 1, p. 17-35, Apr. 1977.

LANCASTER, F. W.; ELLIKER, C.; CONNELL, T. H. Subject analysis. *Annual Review of Information Science and Technology (ARIST)*, [s. l.], v. 24, p. 35-84, 1989.

LIMA, G. Â. B. de O. Representação, recuperação e acesso da informação: a evolução da biblioteca 1.0 à biblioteca 3.0. In: CAMPOS, Maria Luiza de Almeida; MARCONDES, Carlos Henrique; SOUZA, Joice Cleide Cardoso Ennes de; RODRIGUES, Ana Célia; VOGEL, Michely Jabala Mamede; OLIVEIRA, Lucia Maria Velloso de (org.). *Produção, tratamento, disseminação e uso de recursos informacionais heterogêneos: diálogos interdisciplinares*. Niterói: IACS/UFF, 2018. p. 80-88.

LIMA, G. Â. B. de O.; MACULAN, B. C. M. dos S. Análise de assunto a partir de uma perspectiva histórica do ARIST. *Ciência da Informação*, Brasília, v. 41, n. 1, p. 22-35, jan./abr. 2014.

- MEDEIROS, A. L. As bibliotecas na antiguidade. *Memória e Informação*, Rio de Janeiro, v. 3, n. 2, p. 69-85, jul./dez. 2019.
- MONTEIRO, S. D.; GIRALDES, M. J. C. Aspectos lógico-filosóficos da organização do conhecimento na esfera da Ciência da Informação. *Informação & Sociedade*, João Pessoa, v. 18, n. 3, p. 13-27, set./dez. 2008.
- NAVES, M. M. L. Estudo de fatores interferentes no processo de análise de assunto. *Perspectivas em Ciência da Informação*, Belo Horizonte, v. 6, n. 2, p. 189-203, jul./dez. 2001.
- OLIVEIRA, L. P. de; GRÁCIO, M. C. C.; MARTÍNEZ-ÁVILA, D. A expressão “Tratamento Temático da Informação” em artigos de periódicos nacionais: análise da ocorrência e de suas variantes designativas. *AtoZ: Novas Práticas em Informação e Conhecimento*, Curitiba, v. 9, n. 2, p. 44-56, jul./dez. 2020.
- OLIVEIRA, L. P. de; MARTÍNEZ-ÁVILA, D. Tratamento temático da informação: análise sobre abordagens e enfoques em artigos científicos da área de Ciência da Informação. *Informação em Pauta*, Fortaleza, v. 4, n. 2, p. 83-100, jul./dez. 2019.
- OLIVEIRA, L. P. de; SOUZA, M. A. R. de. A desinformação como pilar da intersecção entre letramento informacional e tratamento temático da informação. *Liinc em Revista*, Rio de Janeiro, v. 17, n. 1, p. 1-19, maio 2021.
- ORTEGA, C. D. Fundamentos da organização da informação frente à produção de documentos. *TransInformação*, Campinas, v. 20, n. 1, p. 7-15, jan./abr. 2008.
- PANDO, D. A. *Epistemologia da organização da informação: uma análise de sua cientificidade no contexto brasileiro*. 2018. Tese (Doutorado em Ciência da Informação) –

Faculdade de Filosofia e Ciências, Universidade Estadual Paulista Júlio de Mesquita Filho, Marília, 2018.

REDIGOLO, F. M. O processo de análise de assunto na catalogação em bibliotecas universitárias para a proposta de normalização: estudo de observação com protocolo verbal. *Brazilian Journal of Information Science: Research Trends*, Marília, v. 8, n. 1/2, p. 1-24, 2014.

SALES, R. de. *A presença de Kaiser no quadro teórico do tratamento temático da informação (TTI)*. 2012. Tese (Doutorado em Ciência da Informação) – Faculdade de Filosofia e Ciências, Universidade Estadual Paulista Júlio de Mesquita Filho, Marília, 2012.

SANTOS, R. F. dos. Indexação em repositórios digitais: uma abordagem sobre o metadado assunto da Biblioteca Digital de Monografias da UFRN. *Revista Informação na Sociedade Contemporânea*, Natal, n. esp. 2017.

SIMÕES, M. da G. de M. Do conceito e dos princípios metodológicos de indexação: subsídios para uma análise crítica. In: JORENTE, Maria José Vicentini; LLANES PADRÓN, Dunia (org.). *Uma mirada a la Ciencia de la Información desde los nuevos contextos paradigmáticos de la posmodernidad*. Marília: Oficina Universitária; São Paulo: Cultura Acadêmica, 2017. p. 151-176.

SOUSA, B. P. de. O assunto, a representação: uma análise com foco na variação terminológica a partir da Série Estudos Avançados em Organização do Conhecimento (ISKO Brasil). In: BARROS, Thiago Henrique Bragato; LAIPELT, Rita do Carmo Ferreira (org.). *Organização e representação do conhecimento em múltiplas abordagens*. São Paulo: Pimenta Cultural, 2022. p. 44-93.

SOUSA, B. P. de. Representação temática da informação documentária e sua contextualização em biblioteca. *Revista*

Brasileira de Biblioteconomia e Documentação, São Paulo, v. 9, n. 2, p. 132-146, jul./dez. 2013.

SOUSA, B. P. de; FUJITA, M. S. L. A classificação bibliográfica no contexto do tratamento temático da informação: um estudo com protocolo verbal individual em bibliotecas do Instituto Federal de Educação, Ciência e Tecnologia (IF's). *Revista ACB: Biblioteconomia em Santa Catarina*, Florianópolis, v. 18, n. 1, p. 796-813, jan./jun. 2013.

WEITZEL, S. da R.; SANTOS, A. R. dos. Coleções especiais em bibliotecas universitárias: desafios para a nossa geração. *In: CAMPOS, Maria Luiza de Almeida; MARCONDES, Carlos Henrique; SOUZA, Joice Cleide Cardoso Ennes de; RODRIGUES, Ana Célia; VOGEL, Michely Jabala Mamede; OLIVEIRA, Lucia Maria Velloso de (org.). Produção, tratamento, disseminação e uso de recursos informacionais heterogêneos: diálogos interdisciplinares*. Niterói: IACS/UFF, 2018. p. 61-70.

CAPÍTULO 9

A PRÁTICA DE INDEXAÇÃO EM CORRESPONDÊNCIAS: PROPOSTA DE MODELO DE LEITURA E REPRESENTAÇÃO DOCUMENTÁRIA

Amanda Freitas dos Santos

Márcia Ivo Braz

1 INTRODUÇÃO

No processo de organização da informação, a representação de assuntos caracteriza uma atividade essencial, que deve considerar as características dos tipos documentais que compõem o acervo, visando uma melhor indexação.

Centros de documentação apresentam uma grande variedade de arquivos, incluindo documentos privados pertencentes à atividade civil de personalidades que, de alguma maneira, deixaram seu legado histórico em diversas áreas de conhecimento e na sociedade. Dentre os arquivos privados, as correspondências representam um tipo documental que apresenta características peculiares, e é fundamental que sejam empregados métodos específicos de leitura e representação de seu conteúdo.

Desta maneira, chegamos à seguinte indagação: como promover leituras e representações adequadas dos conteúdos encontrados nas correspondências?

Nosso objetivo é identificar um possível modelo de leitura e representação documental para a indexação de correspondências. Para tanto, foi realizada uma revisão de literatura acerca dos seguintes temas: representação da informação, indexação e tipologia documental correspondência. Em seguida, foi construído um quadro modelo para leitura e síntese documental de correspondências, aplicado em exemplos extraídos do acervo virtual da Fundação Joaquim Nabuco.

A indexação é uma atividade fundamental, pois permite a representação informacional adequada dos diversos documentos em bibliotecas ou arquivos, em formato impresso, digitalizado ou documentos que já nascem e cumprem seu ciclo em formato eletrônico. Porém representa um grande desafio para os profissionais da informação a diversidade dos suportes e sua tipologia, já que solicitam análises para identificação do seu teor.

Entre as tipologias documentais, os conjuntos de correspondências possuem importante valor histórico e social, representando um patrimônio cultural, disponível como fonte de pesquisa e informação com significância regional e nacional, logo, é fundamental que se considere a importância da análise de assuntos no tratamento e na representação da informação documental.

Propor um modelo de leitura e representação documentária pode ser uma referência para os indexadores, pois permite um parâmetro na leitura de correspondências e na escolha dos descritores, auxiliando a facilitação e adequação da descrição de sua tipologia documental.

2 A PRÁTICA DA INDEXAÇÃO

Araújo Junior (2007, p. 23) define “a indexação como processo de representação do conteúdo dos documentos, sendo um elemento fundamental para o processo de busca e recuperação da informação”. Não diferente, Cintra (1983) considera a indexação como uma operação de tradução de documentos em termos documentais, onde são extraídos descritores, cabeçalhos de assuntos e termos chaves, que tem como função, expressar as ideias contidas no documento.

Guinchat e Menou (1994, p. 177), também descrevem algumas etapas no processo de indexação: Tomar conhecimento sobre o documento; Ler as principais partes do documento, para assimilação e compreensão de seu tema, extraíndo do mesmo, as ideias mais importantes; Selecionar os assuntos mais relevantes e descrevê-los em palavras ou termos; Verificar a pertinência dos termos ou palavras selecionados, observando a sua adequação à temática e conceitos contidos no documento; Traduzir os termos selecionados em descritores da Linguagem Documentária utilizada pela unidade de informação, caso a mesma utilize; Formalizar a descrição, se houver regras de apresentação ou escrita, estabelecida pela Instituição.

O exame de um documento tem início a partir da leitura do texto, o que influencia na qualidade da sua análise, síntese e representação. Fujita (2003) evidencia que na perspectiva de tratamento informacional, a indexação está atrelada às operações de tratamento temático, sendo a leitura documentária a principal etapa.

Sousa e Fujita (2014, p. 28) também consideram que:

A primeira etapa da indexação decorre da análise de assunto, que é realizada por meio da leitura documentária do bibliotecário, denominado também de indexador. Por meio desse tipo de leitura se compreende de maneira geral o documento, para assim identificar e selecionar os termos e posteriormente realizar a representação do assunto, a fim de disponibilizá-lo para recuperação no catálogo.

Na análise de assuntos, Fujita (2003, p. 64), subdivide o processo de leitura em três estágios que correspondem a “compreensão do conteúdo do documento; identificação dos conceitos que representam este conteúdo e seleção dos conceitos válidos para recuperação”.

A leitura documentária e a representação documental são etapas onde pode haver dificuldades para o indexador, pois conforme o tipo documental pode haver questões onde “o leitor indexador apresenta dificuldades quanto à identificação e seleção de conceitos representativos do assunto do documento” (Silva; Fujita, 2004, p. 135). Existem vários fatores que levam à dificuldade do indexador, dentre eles a adequação das leituras e representações conforme as diversas tipologias documentais de registros do conhecimento,

demonstrando a necessidade de adoção de modelos e políticas que direcionem os procedimentos do indexador, de preferência descritos em documentos institucionais, como as políticas de indexação.

Diante da interdisciplinaridade e do caráter fluido da expressão da Ciência de uma forma geral, a informação encontra-se estruturada em diferentes formas, a exemplo dos impressos, manuscritos, imagens, multimídias e eletrônicos, o que representa um desafio para a representação e organização da informação.

3 TIPOLOGIAS DOCUMENTAIS: CORRESPONDÊNCIAS

Identificadas como documentos específicos, as correspondências apresentam características gerais comuns, produzidas e utilizadas basicamente para comunicação. Porém, esta tipologia documental é encontrada estruturalmente de maneira diversificada, subdivida em documentos elaborados tanto para fins pessoais, como para fins de caráter oficiais, a exemplo de cartas, telegramas, bilhetes, e-mails, ofícios, cartas-convites, entre outros, que são correspondências utilizadas como meio de comunicação pessoal ou entre autoridades privadas e públicas.

De acordo com a Legislação Postal Internacional, Camargo (2011, p. 60) descreve correspondências como um “objeto que contém comunicação ou nota atual e pessoal, dirigida a outrem”. Desta maneira, as correspondências podem ser consideradas como formas de comunicação produzidas a

pessoa física, jurídica, ou como meios de comunicação entre servidores e órgãos de uma instituição. Essas comunicações podem ser particulares (para cumprir interesses pessoais) ou oficiais (atreladas aos assuntos relacionados às atividades internas de uma organização).

As correspondências são destinadas para fins internos e externos: para as necessidades de comunicação interna entre os departamentos de uma instituição, são utilizados memorandos, despachos e circulares; já em atividades direcionadas às comunicações externas, são utilizados cartas, telegramas, ofícios.

No Manual de Redação de Correspondências e Atos Oficiais, a Fundação Nacional de Saúde (2006, p. 11) inclui as correspondências oficiais nas redações oficiais e define estas como “meio utilizado para o estabelecimento de relações de serviço na administração pública e corresponde ao modo uniforme de redigir atos normativos e comunicações oficiais”.

Para sintetizar as definições, baseado na definição de carta de Camargo (2011) e no Manual de Redação da Presidência da República, Brasil (2018), apresentaremos no Quadro 1 algumas definições de tipos de correspondências:

Quadro 1: Tipos de correspondências

Correspondência	Definições e características
Carta	Documento em papel, com ou sem envoltório, com comunicação ou nota atual e pessoal.
Telegrama	Toda comunicação pessoal ou oficial, utilizada para mensagens especiais ou urgentes e expedida por meio de telegrafia, telex.
Ofícios	Comunicação oficial expedida entre autoridades
Aviso	Comunicação oficial similar ao ofício, porém somente expedida exclusivamente por Ministros de Estados, para autoridades de mesma hierarquia.
Memorando	Comunicação interna entre Unidades administrativas em um Órgão, com igual ou diferente nível hierárquico.
Correio eletrônico (E-mail)	Comunicação para transmissão de mensagens ou documentos, realizados via internet.
Fax (fac-simile)	Comunicação utilizada para transmissão de mensagens urgentes e para envio de documentos, realizados por via telefônica.

Fonte: As autoras.

Para análise estrutural, basicamente as correspondências, em sua maioria, apresentam endereçamento, signatários, destinatário, assunto principal, saudações, assinaturas e período de envio. Já nas correspondências utilizadas em comunicações oficiais devem basicamente apresentar impessoalidade, pronomes de tratamentos, fechos de saudações, Identificação e função do signatário, endereçamento, destinatário, periodização, logomarcas e assinaturas institucionais.

Na Figura 1, a seguir, temos como um dos exemplos de correspondência oficial, a estrutura básica de um ofício:

Figura 1: Modelo de ofício

<p>FRENTE</p> <p>MINISTÉRIO DA SAÚDE FUNDAÇÃO NACIONAL DE SAÚDE (LOGOMARCA) Excelentíssimo Senhor Presidente da República NOME (em caixa alta) Palácio do Planalto 70.150-900 – Brasília-DF</p>
<p>VERSO</p> <p>Remetente: NOME (em caixa alta) Cargo (em caixa alta e baixa) Setor de Autarquias Sul Quadra 4 - Bloco N 70.070-0400 – Brasília-DF</p>

Fonte: Brasil (2006, p. 22).

Nos manuais de comunicação oficial figuram outros modelos de correspondência nas quais também podem ser identificadas as informações ligadas a quem envia e quem recebe, o assunto, o local e quando foi enviado.

Deve-se salientar que apesar destas informações serem elementos padrão na elaboração e estruturação de correspondências, há documentos que apresentam lacunas sobre as informações básicas. Isto acontece porque na produção das correspondências, em sua fase pessoal ou

corrente administrativa, a elaboração padronizada não é observada ou adotada, não apresentando algum componente que seria importante para a sua representação.

Tanto as correspondências produzidas para fins administrativos, como as produzidas para fins pessoais são consideradas como documentação de arquivo e podem ganhar importância à administração interna, quando está em fase corrente ou intermediária, ou à história da instituição, quando se encontram na fase de documentos permanentes.

Em unidades de documentação, as correspondências pessoais ou institucionais ligadas a diversas temáticas, períodos e personalidades, podem representar um conjunto de documentos de valor histórico, social, políticas e cultural, pelo qual o pesquisador se apropria, os transformando em fontes de pesquisa e memória.

4 INDEXAÇÃO DE CORRESPONDÊNCIAS: PROPOSIÇÃO DE MODELO DE LEITURA E REPRESENTAÇÃO DOCUMENTÁRIA

A ABNT 12676 (Associação Brasileira de Normas Técnicas, 1992) apresenta e indica três estágios de que devem ser levados em consideração na prática de indexação: o exame dos documentos e estabelecimento de seu assunto; a identificação dos conceitos presentes no assunto e, por fim, a tradução dos conceitos em termos que representam uma linguagem de indexação, de acordo com a linguagem documentária utilizada pela Unidade.

Guinchat e Menou (1994, p. 179) afirmam que

a indexação deve responder às questões que o usuário faz para saber qual é o objetivo de um documento e em que este documento pode servi-lo. Pode-se lembrar destes objetivos sistematicamente pelas seguintes perguntas: “O que”? De que forma? Como? Quando? e Onde?

Da mesma maneira, em complemento e auxílio na leitura documental e identificação de conceitos, Fujita (2003) faz referência ao mesmo procedimento baseado em questionamento, onde na leitura, análise e identificação de conceitos é recomendada uma a interrogação textual, onde cada questão será correspondida a um conceito com função particular: Ação -> O que aconteceu? Agente da Ação -> Quem / o que fez isso? Objeto da Ação -> A quem aconteceu? Local -> Onde Aconteceu? Tempo -> Quando aconteceu?

No processo de indexação de correspondências, estes questionamentos podem auxiliar na prática de leitura e representação da tipologia documental, pois este tipo documental apresenta características próprias, onde em análise estrutural, basicamente as correspondências em sua maioria, devem apresentar endereçamento, signatários, destinatários assunto principal e período de envio.

Camargo (2011, p.18) faz as pertinentes indagações de pesquisa perante o objeto documental de correspondências: “Que é a correspondência? Quem escreve? Quem lê? O que lê? Em que época? Por que motivos?”.

Estes questionamentos são próprios às respostas estruturais das correspondências e devem ser adotados como

parâmetros para o auxílio na leitura e representação documentária. Assim, é importante que diante dos objetos de análise, sejam respondidas essas principais questões, para que a indexação possa contemplar os principais e importantes tópicos, presentes nas correspondências.

Por conseguinte, para análise, compreensão e representação das correspondências, com base nas etapas de indexação indicados na Norma 12676 (ABNT, 1992) e por Guinchat e Menou (1994), propomos os seguintes parâmetros:

Quadro 2: Parâmetros para modelo de Indexação

PARÂMETROS PARA INDEXAÇÃO DE CORRESPONDÊNCIAS	
1	Tomar conhecimento sobre a correspondência;
2	Leitura das principais partes da Correspondência;
3	Elaboração de resumo;
4	Seleção de descritores, baseado nos questionamentos;
5	Tradução dos descritores de acordo com a LD adotada na Instituição.

Fonte: As autoras.

Nestas etapas, como auxílio, é importante a elaboração de Quadros que serão utilizados para promover e facilitar o processo de indexação face aos questionamentos e respostas elaboradas no processo da verificação e representação das correspondências. Desta maneira, como proposição para um modelo de leitura e representação, baseada nos questionamentos de Camargo (2011) e Fujita (2003), adota-se para análise de correspondências os seguintes questionamentos: Quem escreve? Quem lê? O que lê? Onde? Quando?

Quadro 3: Exemplo de quadro auxiliar

CORRESPONDÊNCIA (TIPO E CLASSIFICAÇÃO)	
QUEM ESCREVE?	
QUEM LÊ?	
O QUE LÊ?	
ONDE?	
QUANDO?	

Fonte: As autoras.

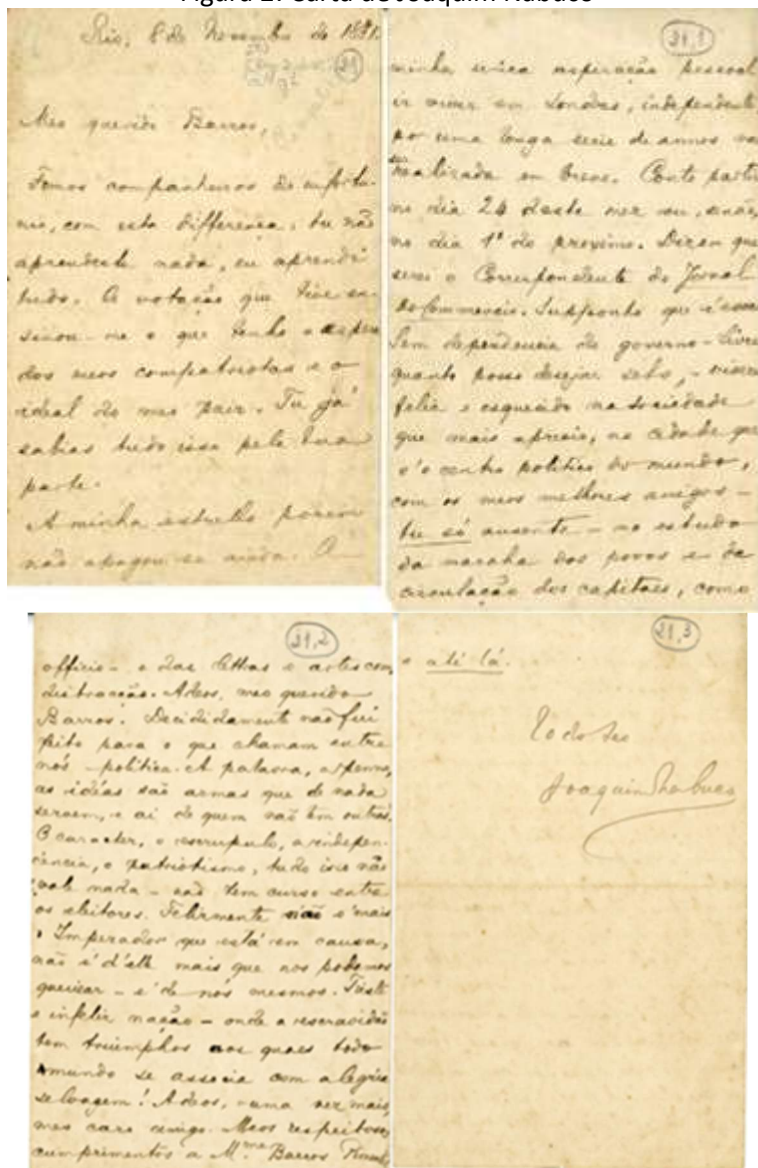
Deve-se salientar, contudo, que durante o processo de indexação algumas cartas podem não responder exatamente a todos os questionamentos, podendo haver lacunas por falta de adoção da estrutura oficial das correspondências, mas que sejam respondidas as questões possivelmente identificadas. É também oportuno que a quantidade de descritores utilizados para representação deve ser adequada às necessidades e objetivos da unidade de informação.

Neste sentido, serão apresentados, a seguir, alguns exemplos de modelos de leitura e representação de correspondências, cujos documentos estão disponíveis no catálogo da Fundação Joaquim Nabuco.

A descrição de conteúdo através da indexação, conforme considerações anteriores, deve ser realizada pela instituição. Na identificação dos campos é possível verificar a síntese documental, auxiliando na escolha dos descritores, conforme a política institucional e a linguagem documentária adotada.

A primeira correspondência analisada, será a Figura 2:

Figura 2: Carta de Joaquim Nabuco



Fonte: Fundação Joaquim Nabuco (2021).

JOAQUIM NABUCO: JN Cap2 doc. 21,3 (CARTA)	
QUEM ESCREVE?	Joaquim Nabuco
QUEM LÊ?	Barros Pimentel
O QUE LÊ?	Política brasileira. Viagem à Londres.
ONDE?	Rio de Janeiro
QUANDO?	1891

Resumo: Carta de Joaquim Nabuco ao seu amigo Barros Pimentel. Fala sobre a política e suas decepções com os compatriotas e com o ideal do país, da triste e infeliz Nação. Avisa sobre sua viagem para viver em Londres e manda despedidas e cumprimentos.

Na Figura a seguir, temos a tipologia telegrama:

Figura 3: Telegrama Para Edson Nery da Fonseca



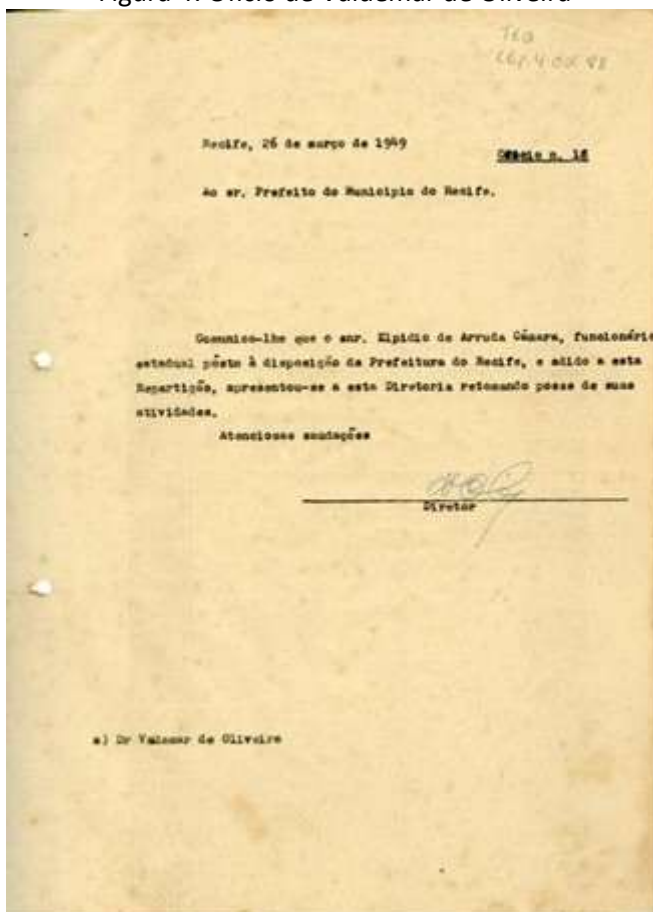
Fonte: Fundação Joaquim Nabuco (2021).

Resumo: Telegrama de José Carlos Poroca a Edson Nery da Fonseca. Envia cumprimentos pelo recebimento de ato de

condecoração com a Medalha dos Guararapes, por reconhecimento de trabalho em prol da Cultura Nacional.

No exemplo seguinte, representado pela Figura 4, temos a tipologia documental ofício.

Figura 4: Ofício de Valdemar de Oliveira



Fonte: Fundação Joaquim Nabuco (2021).

TEATRO: TEa CEp4 doc. 88 (OFÍCIO)	
QUEM ESCREVE?	Valdemar de Oliveira
QUEM LÊ?	Manoel César de Moraes Rêgo
O QUE LÊ?	Teatro Santa Isabel. Posse de funcionário.
ONDE?	Recife
QUANDO?	1949

Resumo: Ofício de Valdemar de Oliveira, diretor do Teatro Santa Isabel, a Manoel César de Moraes Rêgo, prefeito do Recife. Comunica a apresentação e posse nas atividades no teatro de Elpidio de Arruda Câmara, funcionário estadual.

Após analisar os exemplos de correspondência, foi possível evidenciar que a instituição de um modelo de leitura apropriado a tipologia documental correspondência e suas variações, o que contribui para a consistência na indexação e representação de conteúdo. Estabelecer um modelo de análise para os tipos documentais que compõem o acervo da unidade de informação, portanto, pode ser uma alternativa interessante para o melhor trabalho dos indexadores e, conseqüentemente, para a recuperação da informação por parte do público usuário.

5 CONSIDERAÇÕES FINAIS

Este trabalho discutiu a representação documental e apresentou um modelo de leitura documentária aplicado às correspondências. A diversidade dos suportes e sua tipologia informacional demandam análises e métodos adequados na indexação dos documentos, o que se torna um desafio ao indexador. Considerar as peculiaridades das tipologias documentais, permite aos centros de documentação, uma melhor representação e disponibilização de seus acervos à

sociedade, pois possibilita uma maior recuperação da informação contida nos documentos.

Com a popularização dos documentos eletrônicos, desde sua criação e durante todo o seu ciclo de vida, bem como a digitalização de acervos, temos uma massa documental bastante considerável e passível de aplicação de metodologias de representação direcionadas a este contexto. Por esse motivo, faz-se necessária a adoção de critérios específicos de análise, políticas de indexação bem estabelecidas e controle terminológico conforme as especificidades das coleções e do contexto de uso.

A adoção de um modelo é importante para nortear a atividade dos diferentes indexadores, face às diversas tipologias documentais. Contudo, deve-se levar em consideração que este modelo pode e deve se adequar a realidade e necessidades internas de cada Instituição e suas diferentes bases de dados.

Assim, a adoção de um modelo de leitura e análise documental, considerando suas tipologias e características peculiares, atrelado à adoção de políticas de indexação e adoção de um vocabulário controlado, pode auxiliar e facilitar a prática da indexação, pois permite uma padronização das atividades internas nas unidades de informação, objetivando dirimir possíveis ruídos e melhorar a precisão na recuperação da informação.

REFERÊNCIAS

- ARAÚJO JUNIOR, R. H. *Precisão no processo de busca e recuperação da informação*. Brasília: Thesaurus, 2007.
- ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. *NBR 12676: Método para análise de documentos: determinação de assuntos e seleção de termos de indexação*. Rio de Janeiro: ABNT, 1992. Disponível em: <https://www.abntcatalogo.com.br/norma.aspx?ID=3163> Acesso em: 21 jun. 2021.
- BRASIL. Presidência da República. *Manual de Redação da Presidência da República*. 3. ed. Brasília: Presidência da República, 2018. Disponível em: <http://www4.planalto.gov.br/centrodeestudos/assuntos/manual-de-redacao-da-presidencia-da-republica/manual-de-redacao.pdf>. Acesso em: 20 abr. 2021.
- CAMARGO, M. R. R. M. *Cartas e escritas: práticas culturais linguagem e tessitura da amizade*. São Paulo: Editora Unesp, 2011.
- CINTRA, A. M. M. Elementos de linguística para estudos de indexação. *Ciência da Informação*, Brasília, v. 12, n. 1, p. 5-22, 1983. Disponível em: <https://brapci.inf.br/index.php/res/v/22075>. Acesso em: 21 jun. 2021.
- FUJITA, M. S. L. A identificação de conceitos no processo de análise de assunto para indexação. *Revista Digital de Biblioteconomia & Ciência da Informação*, Campinas, v. 1, n. 1, p. 60-90, 2003. DOI: 10.20396/rdbci.v1i1.2089. Disponível em: <https://periodicos.sbu.unicamp.br/ojs/index.php/rdbci/articloe/view/2089>. Acesso em: 21 jun. 2021.

FUNDAÇÃO NACIONAL DE SAÚDE (Brasil). *Manual de redação de correspondências e atos oficiais*. Brasília: Funasa, 2006. Disponível em: http://www.funasa.gov.br/site/wp-content/files_mf/adm_redOficial.pdf. Acesso em: 10 abr. 2021.

GUINCHART, C.; MENO, M. *Introdução geral as ciências e técnicas da informação e documentação*. 2. ed. Brasília: IBICT, 1994.

SILVA, M. R.; FUJITA, M. S. L. A prática de indexação: análise da evolução de tendências teóricas e metodológicas. *Transinformação*, Campinas, v. 16, n. 2, p. 133-161, 2004. Disponível em: <http://hdl.handle.net/11449/114455>. Acesso em: 29 abr. 2021.

SOUSA, B. P.; FUJITA, M. S. L. A análise de assunto no processo de indexação: um percurso entre teoria e norma. *Informação & Sociedade: Estudos*, João Pessoa, v. 24, n. 1, 2014. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/92987>. Acesso em: 21 jun. 2021.

CAPÍTULO 10

UMA CONTRIBUIÇÃO PARA REPRESENTAÇÃO DESCRITIVA DE ARTEFATOS TRIDIMENSIONAIS DA CULTURA POPULAR

Karin Herculano Picado

Edilene Toscano Galdino dos Santos

1 INTRODUÇÃO

A sociedade contemporânea possui um vasto legado para a acepção do que vem a ser cultura e cultura popular, contudo diante deste universo conceitual, não vamos nos prender ao ponto de vista filosófico ou ontológico dessa terminologia.

Faz-se uma rápida passagem pelo entendimento de cultura para que se possa compreender o processo descritivo do ponto de vista de importância do artefato tridimensional como elemento para a formação de memória cultural de um povo. Assim, foi construído este escrito, em que foca a biblioteca como um espaço também para a memória de um povo e a partir de um fazer técnico do bibliotecário que está na montagem de uma proposta prática de como representar e descrever artefatos da cultura popular no contexto do uso de conceitos biblioteconômicos.

E nesse sentido, uma pesquisa no âmbito da área de Representação Descritiva, se transforma em proposta para descrição de artefatos tridimensionais da cultura popular,

sendo este estudo uma contribuição para a área da descrição, tendo em vista que aborda elementos que são de pouca visibilidade na área biblioteconômica como os artefatos da cultura popular.

Assim, diante de uma necessidade do fazer bibliotecário, direciona-se o estudo a compreensão e criação de uma descrição para os artefatos tridimensionais da cultura popular, contribuindo para área da Representação Descritiva que abrange os objetos tridimensionais e realia.

2 CULTURA: LEGADO COMUM DE TODA A HUMANIDADE

Compreendemos que conservar e preservar o patrimônio histórico, cultural e social com o propósito de perpetuar as diversas possibilidades de construção de memória, estão entre as responsabilidades das bibliotecas, em harmonia com a declaração preparada pelo International Federation of Library Associations and Institutions (IFLA).

Segundo Santos (1987), a cultura é um produto da história coletiva, que por suas transformações e seus benefícios, as forças sociais se defrontam. Isso nos parece particularmente importante quando consideramos as “mazelas culturais” de um povo como, por exemplo: o analfabetismo; o controle do conhecimento e seus benefícios por uma pequena elite; a pobreza do serviço público de educação e de formação intelectual das novas gerações. Santos (1987), considera ainda que não há nada mais incisivo como marcador da identidade de um povo do que a cultura em

que este povo se define. Porém cultura, para este autor, seria uma preocupação contemporânea. Preocupação esta, em entender os muitos caminhos que conduziram os grupos humanos às suas relações presentes e suas perspectivas de futuro.

Sendo assim, compreendemos que a discussão de cultura sempre remete ao processo, à experiência histórica, não fazendo sentido ver a cultura como um sistema fechado. Então, cultura diz respeito à humanidade como um todo e ao mesmo tempo, diz respeito a cada um dos povos, nações, sociedades e grupos humanos. E sem ela, tal identidade desaparece, dissolve-se, deixando ali, como consequência, uma amorfa massa humana. Por isso, nos juntamos a ele ao considerarmos imperativo o reconhecimento, incentivo e preservação da cultura popular de forma que as gerações futuras se sintam não só herdeiras, mas também valedouras desse precioso tesouro.

Apesar de sua reconhecida importância, a Arte Popular Brasileira, Santos (1987) considera que ela seja em grande parte desconhecida e subvalorizada em nosso próprio país.

Produzido por pessoas nascidas nos meios menos privilegiados economicamente da população e que, na sua maior parte, não tiveram acesso à escolaridade formal, esse acervo, a despeito de suas qualidades intrinsecamente formais e estéticas desempenha um papel importante na preservação da memória da produção artística desse segmento.

3 ASPECTOS HISTÓRICOS DA REPRESENTAÇÃO DESCRITIVA

Na evolução dos tempos já havia por parte da humanidade tentativas de registrar o conhecimento como legado ancestral. Do ponto de vista da Biblioteconomia temos a exemplo: Calímacus de Cirene na Biblioteca de Alexandria, no Egito Antigo, procurou meios para recuperar a produção intelectual da época registrada nos rolos de papiro, criando um catálogo denominado Pinakes. Diante do aperfeiçoamento da sociedade, surge a imprensa, o aumento e a disseminação da produção do conhecimento cresceu exponencialmente, o que forçou criações e aperfeiçoamentos de técnicas que viabilizassem a organização do conhecimento produzido para que o mesmo pudesse ser recuperado. (Ribeiro; Chagas; Pinto, 2007).

Accart (2012) diz que Dewey influenciou as bibliotecas a abrirem suas portas e seus acervos para democratizá-las e Paul Otlet sonhou, no final do século XIX, em facilitar o acesso do maior número de pessoas à informação graças a um complexo conjunto de bibliotecas conectadas por canais telegráficos e telefônicos. Paul Otlet junto com Henry La Fontaine, antes da Primeira Guerra, apresentaram ao mundo com diversas organizações para disseminação do conhecimento e assim montar uma rede de conhecimento mundial. (La Fontaine; Otlet, 1895).

Na década de 1960, segundo Mey e Silveira (2009), a Biblioteconomia passou por uma revolução no que tange ao aprimoramento das técnicas que visassem à recuperação da

informação. Com isso, houve a necessidade de utilizar recursos computacionais já disponíveis à época: nos Estados Unidos da América (EUA), o projeto *Machine Readable Cataloging* (MARC), que tinha como propósito ser um recurso computacional que transformasse a catalogação manual realizada em fichas catalográficas, em registros legíveis por máquina por meio da codificação dos registros bibliográficos, transformando-os em catálogos em linha.

Na tentativa de demonstrar a Catalogação como: uma disciplina responsável pelo processo de representação documentária; utilizamos o levantamento histórico na intenção de construir a epistemologia da Catalogação baseada em sua historicidade. Num breve histórico da Catalogação percebe-se a existência de uma preocupação constante na busca por modelos e métodos adequados para a construção de alguma forma de descrição que represente os documentos produzidos e para sua gestão.

Entre os principais objetivos da Representação Descritiva, segundo Santos (2007), por meio do uso de padrões de representação e de descrição do documento, aceitos internacionalmente, definir e homogeneizar dados que sirvam como embasamento para sua recuperação e, desse modo, atender aos usuários de forma eficiente e eficaz, contribuindo para a produção do conhecimento. Portanto, a função da Representação Descritiva seria, segundo essa autora representar o item de modo a mostrar suas características físicas e de conteúdo, garantindo ao usuário, sua localização, identificação e uso.

González de Gomez e Machado (2008) afirmam que existem outros espaços que também abrigam acervos que necessitam ser consultados. Espaços estes, onde são alocados às inscrições do conhecimento em quaisquer meios, que podem ser denominados de bibliotecas, museus, laboratórios etc.

A IFLA instituiu em 2009, a Declaração dos Princípios Internacionais de Catalogação, que estabeleceu uma série de padrões que normatizam os registros bibliográficos e de outros tipos de materiais, tais como: a *International Standard Bibliographic Description* (ISBD) que são normas que apresentam pontuações para cada área da descrição bibliográfica e determinam fundamentos para a construção de catálogos (Mey; Silveira, 2010).

A Declaração de Princípios Internacionais de Catalogação instituída pela IFLA (Declaração..., 2016, p. 11-12) também se refere ao catálogo como o canal de comunicação entre o usuário e o acervo. Assim entendemos que o catálogo deve ser um instrumento efetivo e eficiente que permita ao usuário encontrar recursos bibliográficos em uma coleção como o resultado de uma busca, usando atributos e relações entre recursos.

Nessa perspectiva, observa-se que esses princípios corroboram ao que estabelece a terceira e quarta lei de Ranganathan “A cada livro seu leitor” e “A cada leitor o seu livro”. Dito isso, entende-se que a principal função de um catálogo é permitir que determinado item possa ser encontrado pelo usuário.

Para que cumpra sua função efetiva o catálogo precisa ser dotado de algumas características indispensáveis (Mey; Silveira, 2010, p. 127-128): integridade significa fidelidade, honestidade na representação, transmitindo informações passíveis de verificação; clareza significa que a mensagem deve ser compreensível aos usuários; precisão significa que cada uma das informações só pode representar um único conceito, sem dubiedades ou dúvidas; lógica significa que as informações devem ser organizadas de modo lógico; consistência significa que a mesma solução deve ser sempre usada para informações semelhantes. Assim, para desempenhar essas funções a Catalogação faz uso de semântica, sintaxe e semiótica próprias, conforme Mey e Silveira (2010, p. 128) apontam:

A sintaxe se constitui de posição e pontuação. Ou seja, cada elemento descritivo, ao registrar uma característica ou atributo de uma entidade (no sentido dos FRBR), possui uma posição determinada. Tal posição se encontra explícita no registro bibliográfico elaborado por meio da ISBD. [...] A semiótica dá significado aos termos em sua posição e respectiva pontuação precedente.

Atualmente, no Brasil, o código de catalogação mais utilizado é o *Anglo-American Cataloguing Rules* (AACR2), traduzido para o português em 2002. O AACR2, como é conhecido, é dividido em duas partes: parte 1: referente à descrição bibliográfica, incluindo as áreas da *International Standard Book Description* (ISBD), em 13 capítulos de acordo com o tipo de suporte que se deseja representar. E parte 2: referente aos Pontos de Acesso, Títulos Uniformes e

Remissivas, com seis capítulos, numerados de 21 à 26, bem como seis anexos e o índice.

Para o desenvolvimento do estudo em pauta foi destacado o capítulo 10 do referido Código, referente à Artefatos tridimensionais e Realia que, segundo o AACR2 (Código, 2010) abrange as regras para descrição de artefatos tridimensionais de todos os tipos, como modelos, dioramas, jogos (incluindo quebra-cabeças e simulações), esculturas e outras obras de arte tridimensionais, objetos para exposição, máquinas e vestuários. Objetos que aparecem ao natural.

Mey e Silveira (2010), afirmam que a utilização do código na prática da Catalogação exige a interpretação e uso adequado das regras da catalogação, a hermenêutica e a análise da entidade a ser representada. As áreas e elementos representados, bem como a pontuação irão constituir a configuração do catálogo e sua relação com o usuário.

As características da Catalogação aliadas ao significado, sintaxe, semântica, semiótica, estabelecem juntas os fundamentos necessários para o desenvolvimento eficaz do processo comunicativo entre o usuário e o acervo representado no catálogo, onde o foco será sempre a recuperação relevante da informação, no caso desse estudo, a recuperação do acervo de cultura popular.

4 CATALOGAÇÃO DE ARTEFATOS TRIDIMENSIONAIS E REALIAS

A catalogação (pode ser chamada também de representação bibliográfica ou descrição bibliográfica, mas

cuidado para não confundir com a representação de seu conteúdo, atividade pertinente à indexação e linguagens documentárias) é aqui entendida de acordo com os autores que abordaremos nesta seção, como normas, formatos ou códigos que registram de alguma forma a informação, independente do suporte, que deve possuir algumas características: integridade, clareza, precisão, lógica e consistência.

São considerados conjuntos de informações codificadas para representar cada um dos registros do conhecimento existentes em acervos (Mey; Silveira, 2009). Esses itens visam atender da melhor forma possível à necessidade do usuário.

Catalogar é uma tarefa que requer técnica especializada, exige perfeição na formação dos catálogos de bibliotecas, a fim de atender aos interesses dos usuários, ao fornecer informações concretas, exatas e apropriadas, por meio da identificação do documento solicitado (Souza, 2003 p. 17).

O procedimento básico de catalogação possui algumas etapas, como as identificadas por Guinchat e Menou (1994): tomar conhecimento do documento; Determinar o tipo de documento e as regras aplicáveis a este caso; Determinar o nível de descrição bibliográfica que será utilizado; Identificar, para cada nível de descrição bibliográfica, os dados necessários, na ordem das áreas indicada pela norma ou pelo formato utilizado; Transcrever estes dados de acordo com as regras da norma ou do formato utilizados; Verificar exatidão da descrição e sua conformidade com as normas; Elaborar as

fichas dos catálogos ou outros produtos previstos para o sistema.

O código AACR2 tem como objetivo a normalização da catalogação a nível internacional, subsidiando o tratamento da informação. Utiliza sistema de pontuação e a catalogação pode ser feita pelo suporte físico da obra. Publicado em 1978, com três revisões em língua inglesa entre 1988 e 2002 e alterações em 2003, 2004 e 2005. A tradução de 2002 para o português teve revista à redação e numeração das regras além da inclusão de novos exemplos.

Todas as áreas abrangidas pela AACR2 obedecem a um padrão único – Capítulo 1 do AACR2 – Regras Gerais – com exceção da área 3 (área dos detalhes específicos do material, presente apenas em alguns materiais) e da área 5 (área da descrição física), diferente para cada material descrito.

O capítulo 10 do Código AACR2, referente a Objetos Tridimensionais e Realia, como o próprio título sugere, dita normas para a representação desse tipo de material que, diferente dos livros, são dotados de inúmeros detalhes relevantes para a recuperação da informação. No que concerne a este tópico o AACR2 (Código, 2010, p. 278) afirma que “Quando apropriado, registre os materiais de que são feitos os objetos. Se os materiais não puderem ser mencionados sucintamente, omita-os ou especifique-os em notas”.

O AACR2 ainda recomenda que a fonte principal de informação para a catalogação de objetos tridimensionais seja

o próprio objeto, onde devem ser observados todos os detalhes para o enriquecimento da descrição. (Código, 2010).

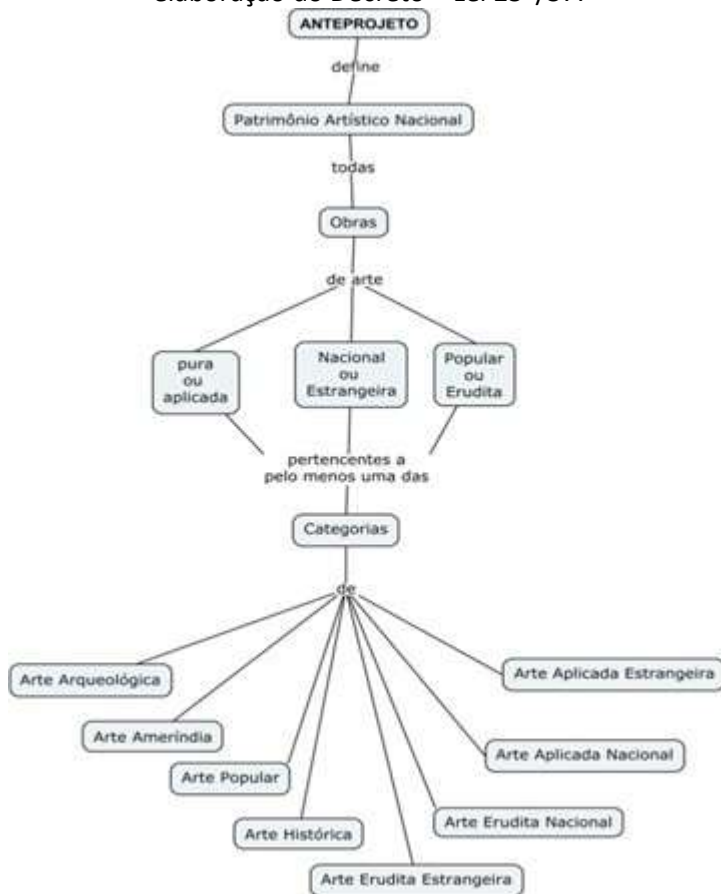
5 UMA CONCEPÇÃO PARA PROTEÇÃO DO PATRIMÔNIO ARTÍSTICO NACIONAL

Em 1936 a pedido de Gustavo Capanema (Ministro da Educação e Saúde), Mário de Andrade (figura pública, intelectual e pesquisador) elaborou um anteprojeto de proteção do patrimônio artístico nacional, que serviu de base para a elaboração do Decreto – Lei 25^o/37.

A concepção do anteprojeto possuía aspectos contemporâneos à sua época e aspectos inovadores: alguns paradigmas atuais de identificação e proteção dos bens que ainda podem ser adotados por estudiosos. Ele é organizado em três capítulos, sendo o primeiro sobre as competências do Serviço do Patrimônio Artístico Nacional (SPAN). No segundo constam os bens culturais que fazem parte do patrimônio artístico nacional e os critérios de seleções desses bens. E o terceiro capítulo dedicado a estabelecer a estrutura interna do SPAN.

De acordo com o anteprojeto de Mário de Andrade, tais bens são classificados em oito categorias, agrupados em quatro livros de tomo: obras de arte pura ou arte aplicada; popular ou erudita; nacional ou estrangeiros, desde que residentes no Brasil.

Figura 1: Mapa Conceitual do Anteprojeto de proteção do patrimônio artístico nacional, que serviu de base para a elaboração do Decreto – Lei 25º/37.



Fonte: Dados da pesquisa (2019).

O anteprojeto de Mário de Andrade foi inovador em sua época porque as convenções de Haia de 1899 e 1907, o Pacto de Roerich de 1935 e os projetos preparados sob o patrocínio das Ligas das Nações, não faziam menções aos bens imateriais. E no referido anteprojeto são identificadas duas tendências contemporâneas relevantes: proteção dos bens

imateriais (e as várias categorias que podem ser incluídas) e a extensa preocupação com o conceito de bem cultural, considerando os merecedores da tutela não somente os monumentos grandiosos, mas também os pequenos artefatos que contribuem para a formação de uma cultura nacional.

6 UMA CONTRIBUIÇÃO PARA REPRESENTAÇÃO DESCRITIVA EM ACERVO DE CULTURA POPULAR

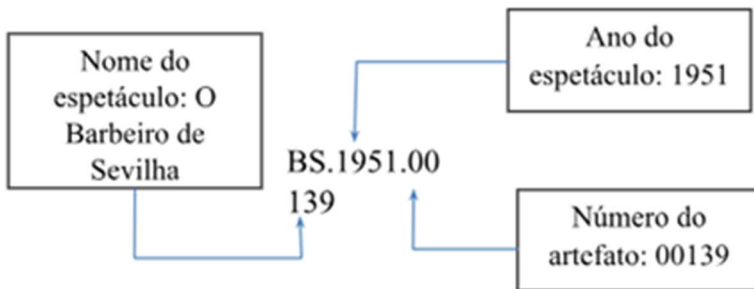
No desenvolvimento de uma proposta para descrição de artefatos tridimensionais da cultura popular considerou os princípios da Catalogação que nortearam a representação desses documentos não tradicionais. Sobre isso, Fusco (2011) afirma que a Catalogação se refere à representação como ‘algo no lugar de’. Para tanto são necessários o uso de padrões internacionais que possibilitem o intercâmbio entre bases de dados de diferentes unidades de informação. Dessa forma, o objetivo é fornecer contribuição para a organização de acervo de peças tridimensionais de coleções especiais de cultura popular que se configuram na representação descritiva do acervo dos artefatos tridimensionais.

Ao analisar o trabalho de Azevedo e Viana (2006), verificamos que esses autores definem como critério para a recuperação do acervo de figurino do Theatro Municipal de São Paulo, o número de localização do vestuário. Na proposta dos referidos autores, o número de localização da vestimenta é composto na seguinte forma:

- Primeiramente duas letras em maiúsculo, associadas às iniciais do espetáculo do qual o figurino pertence;

- Seguido do ano do espetáculo;
- Por último uma sequência composta por cinco dígitos (o número do artefato);
- Cada grupo do código, separados por ponto.

Figura 2: Quadro de exemplo do número de localização de peça do figurino, utilizado para a recuperação do acervo do Theatro Municipal de São Paulo



Fonte: a partir de Azevedo e Viana (2006).

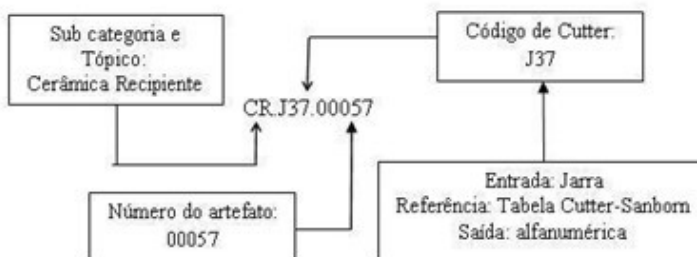
Desta forma, compreende-se que para esses autores o número de localização, da forma como é composto, é relevante para a comunidade que faz uso e manipulação do acervo para sua recuperação eficaz. O trabalho desenvolvido por Gravina (2011) acerca da representação descritiva de objetos de moda contribuiu para o presente trabalho na definição da entrada principal. Verificou-se que para essa autora, a entrada principal é o autor da confecção (criador) da vestimenta, considerado relevante para a recuperação do acervo. Observa-se que a concepção dessa autora está mais próxima do que determina o AACR2 – a entrada principal é o autor do documento.

Entendemos, no entanto, que o número de localização (ou de chamada) da peça de figurino, embora importante para

sua localização, é limitada no sentido de recuperar com precisão a peça.

Na proposta apresentada para o acervo de artefatos tridimensionais de cultura popular, o número de chamada foi composto na seguinte ordem: primeiramente duas letras em maiúsculo, associadas às iniciais da sub-categoria associada ao tópico principal; seguido do Número de Cutter referente ao autor da peça ou da primeira palavra do título ou termo específico, quando não houver autoria; por último uma sequência composta por cinco dígitos (o número de ordem do artefato). Cada grupo do código, separados por ponto.

Figura 3: Quadro de exemplo do número de localização do artefato tridimensional, utilizado para a recuperação do acervo de cultura popular



Fonte: Dados da pesquisa (2019).

Já o trabalho de Brandão (2013), sobre a representação descritiva de instrumentos musicais, colaborou para a adaptação dos campos MARC 21, de modo a hospedar a representação descritiva de objetos tridimensionais, que, no caso deste trabalho são os artefatos tridimensionais de cultura popular.

Dessa maneira, a pesquisa apropriou-se dos elementos indicados pelos trabalhos analisados para a organização do registro bibliográfico, com suas devidas áreas de descrição, buscando-se sua correspondência e/ou equivalência com o MARC 21 e com o AACR2.

O quadro abaixo sintetiza a proposta de representação descritiva de artefatos tridimensionais de cultura popular, compreendendo os elementos indicados pelos trabalhos referenciados e sua correspondência no MARC 21 e no AACR2.

Quadro 1: Proposta de Representação Descritiva Correspondência com os campos MARC 21

MARC 21/AACR2		PROPOSTA
ELEMENTO DE DESCRIÇÃO	CAMPO	
Fonte de aquisição	040	Informações sobre a fonte de aquisição para o item ou suas reproduções.
Número de chamada (ou de localização) do artefato	090	Primeiramente duas letras em maiúsculo, seguido do Número de Cutter referente ao autor da peça ou da primeira palavra do título, quando não houver autoria ou primeira letra do termo específico, quando não houver autoria ou título e uma sequência composta por cinco dígitos (o número do artefato), cada grupo do código, separados por ponto.
Autoria	100	Nome da responsável pela fabricação.
Título uniforme coletivo	243	Título genérico construído pelo catalogador para reunir as

		diversas peças de um mesmo conjunto.
Título principal	245	Título principal completo quando houver (com termos específicos), incluindo informações sobre o material usado na confecção, o subtítulo e a indicação de responsabilidade.
Série	490	Nome da subcategoria (cerâmica, indumentária e fetiche).
Restrição de acesso	506	Utilizada para informações sobre restrições impostas ao acesso aos materiais descritos.
Resumo	520	Informações sobre os detalhes do artefato descrito.
Público alvo	521	Identifica o público específico ou nível para o qual o conteúdo do item descrito é apropriado.
Outros formatos disponíveis	530	Informa a disponibilidade do artefato ser confeccionado, por exemplo, em outro material.
Informações sobre financiamento	536	Informa quem, ou que projeto financiou o artefato.
Condições de uso e reprodução	540	Restringe o direito de reproduzir, exibir, adaptar etc.
Origem	561	Informa a história do material descrito, desde o momento de sua criação até o momento de sua aquisição. Usado para registrar, por exemplo, o nome do doador, ou, o local onde originalmente o artefato se encontrava.
Publicações sobre o material descrito	581	Contém referências de publicações utilizadas para identificar o artefato. Pode ser utilizado também, para registrar

		catálogos de exposição ou de coleção, o que contenha, como fotos etc.
Assunto - nome pessoal	600	Informa sobre um nome pessoal utilizado como assunto atribuído à obra.
Assunto tópico	650	Representado por um termo que represente a categoria, utilizando o Glossário proposto no trabalho.
Assunto - nome geográfico	651	Usado para representar o assunto quando esse for um nome geográfico
Entrada secundária - nome pessoal	700	Nome pessoal como entrada secundária que não tenha sido adotada como entrada principal, por exemplo, colaboradores etc.
Entrada secundária - entidade	710	Nome da entidade como entrada secundária, que não tenha sido adotada como entrada principal, mas que seja considerada ponto de acesso relevante para a recuperação.
Entrada secundária - título uniforme	730	Utilizar para artefatos conhecidos também em língua estrangeira ou para outro uso das peças.
Instituição depositária	850	Informar a instituição responsável pelo acervo do artefato.
Localização ou acesso eletrônico	856	Informações necessárias para dar acesso, por exemplo a imagens do artefato descrito.

Fonte: MARC21 (2000); Dados da pesquisa (2019).

As figuras 4 a 8 mostram as fotos de peças selecionadas para se proceder a catalogação à luz da proposta desenvolvida

nesse trabalho. São apresentadas tanto a catalogação na forma tradicional (em ficha), como na forma padrão MARC 21.

Figura 4: Fotografia da Jarra Cerâmica CR.J37.00057



Fonte: Picado (2019).

A catalogação de um material aqui apresentado para ser descrito utilizando um sistema bibliográfico, o escolhido foi o Biblivre, um *software* livre que se utiliza do MARC21 para sua composição e dado a condição de ser de livre acesso possibilitou realizar os ensaios de descrição para artefatos da cultura popular com o emprego da proposta elaborada como opção de ferramenta para a área dos artefatos tridimensionais.

Figura 5: Print do Formulário Biblivre

Título: Jarra cerâmica em forma de ganso
Localização: CR.J37.00057
Nº do registro: 2

[Editar](#)
[Excluir](#)

Resumo Catalográfico **Formulário** MARC Exemplos

Tipo de material Objeto 3D

Número de chamada - Localização (090) [?]
Classificação CR.J37.00057 **Sa**

Título principal (245) [?]
Gera entrada secundária na ficha Gera entrada para o título **#1**
Número de caracteres a serem desprezados na alfabetação Nenhum caractere a desprezar **#2**
Título principal Jarra cerâmica em forma de ganso **Sa**

Indicação de série (490) [?]
Política de desdobraamento de série Título não desdobraado **#1**
Título da série Cerâmica **Sa**

Notas de público alvo (521) [?]
Notas de público alvo livre para qualquer idade **Sa**

Localização de obras por meio eletrônico (856) [?]
Caminho DigitalMediaController/ **Sd**
Nome do arquivo ?id=MjpEU0NfMDizMV9qYXJyYS5qcGc= **Sf**
URI DigitalMediaController/?id=MjpEU0NfMDizMV9qYXJyYS5qcGc= **Su**
Link em texto jarra cerâmica em forma de ganso **Sy**

Fonte: Biblivre (2019).

Figura 6: Print da catalogação no MARC 21

[Retornar à pesquisa](#) < Anterior 2 / 2 Próximo >

Título: Jarra cerâmica em forma de ganso
Localização: CR.J37.00057
Nº do registro: 2

[Editar](#)
[Excluir](#)

Resumo Catalográfico **Formulário** MARC Exemplos

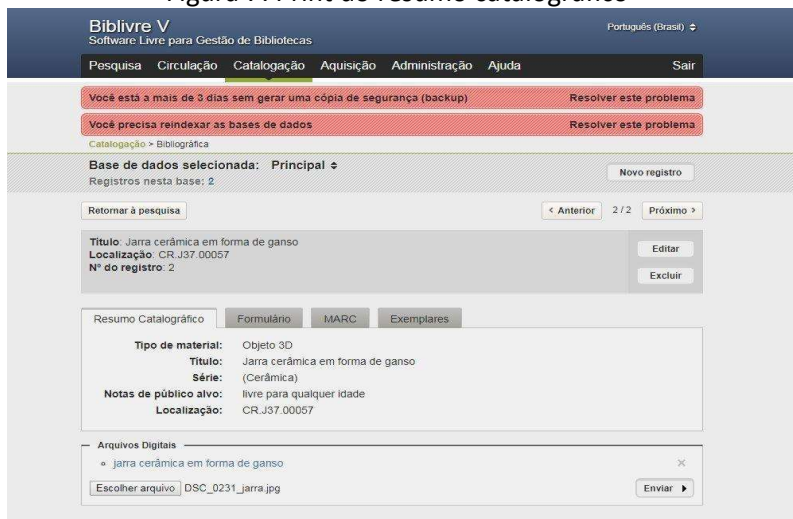
Tipo de material Objeto 3D

```

000 00443nm a2200121 a 4500
001 0000002
005 20190128224659.557
008 190128s||| B|||||por|u
090 __aCR.J37.00057
245 10$aJarra cerâmica em forma de ganso
490 0$aCerâmica
521 __alivre para qualquer idade
856 _d|DigitalMediaController|f?
id=MjpEU0NfMDizMV9qYXJyYS5qcGc=|uDigitalMediaController/?
id=MjpEU0NfMDizMV9qYXJyYS5qcGc=|yjarra cerâmica em forma de ganso
    
```

Fonte: Biblivre (2019).

Figura 7: Print do resumo catalográfico



Fonte: Biblivre (2019).

Figura 8: Print da etiqueta do artefato catalogado



Fonte: Biblivre (2019).

Assim, temos uma aplicação da descrição para artefatos tridimensionais voltado a uma área pouco explorada no seguimento da catalogação brasileira que está voltada para resolver as questões de aplicabilidade de suas orientações disponíveis na teoria da técnica apresentada em códigos e padrões da representação descritiva.

7 CONSIDERAÇÕES FINAIS

O objetivo da presente pesquisa buscou elaborar e definir uma proposta de representação descritiva para seguimento de artefato de cultura popular, com base em padrões internacionais consagrados na literatura, visando à disponibilização desse acervo em um Sistema de Recuperação da Informação (SRI), para propiciar e otimizar sua recuperação e uso.

Quanto à representação descritiva, observamos que apesar dos padrões existentes contemplarem, de um modo geral, os objetos tridimensionais, no qual o artefato de cultura popular se enquadra, são bastante escassos, sobretudo, no que tange as diversas especificidades da cultura popular e de cada elemento que compõe a ampla gama de objetos tridimensionais.

Ademais, como são poucos os trabalhos sobre representação de artefatos de cultura popular na área de Biblioteconomia, para este trabalho, nos baseamos em itens descritos de outras áreas de artefatos tridimensionais.

Assim, nota-se que os itens de cultura popular são extremamente peculiares e não estão expressamente contemplados no capítulo 10 do AACR2 e no MARC21, daí as adaptações que foram propostas para a descrição aqui apresentada.

Procuramos categorizar os artefatos como documento, de acordo com os teóricos da área, e assim, evidenciar a importância da prática de representação descritiva desse tipo de material. Na literatura sobre cultura popular, o artefato

também é considerado documento justificado esse fato, no argumento de que é uma forma individual de cultura material, isto é: manifestações da arte em objetos.

Sendo assim, na representação descritiva de acervo de cultura popular teve como principal desafio adaptar as normas Biblioteconômicas, muito utilizadas para livros e outros documentos em papel ou virtual, de modo a utilizar metadados existentes, que compreendessem a especificidade exigidas para essa tipologia de material tridimensional.

REFERÊNCIAS

ACCART, J-P. *Serviço de referência: do presencial ao virtual*. Brasília: Briquet de Lemos / Livros, 2012.

AZEVEDO, E. R.; VIANA, F. *Breve manual de conservação de trajes teatrais*. São Paulo: Gráfica InPrima, 2006.

BRANDÃO, D. C. *Representação documentária de instrumentos musicais: contribuição para a organização do Museu Instrumental Delgado de Carvalho da Escola de Música da Universidade Federal do Rio de Janeiro (UFRJ)*. 2013. Trabalho de Conclusão de Curso (Especialização em Políticas de Informação e Organização do Conhecimento) – Universidade Federal do Rio de Janeiro / Arquivo Nacional, Rio de Janeiro, 2013.

BIBLIVRE FOR WINDOWS 98. Versão 5.1. 2018. Disponível em: <https://www.biblivre.org.br/index.php/sobre-biblivre/o-programa>. Acesso em: 20 out. 2018.

CÓDIGO DE CATALOGAÇÃO ANGLO-AMERICANO. 2. ed., revisão 2002, reimpressão. São Paulo: FEBAB, 2010.

IFLA. Declaração *dos Princípios Internacionais de Catalogação (PIC)*. Haia: IFLA, 2016. Disponível em: https://www.ifla.org/wp-content/uploads/2019/05/assets/cataloguing/icp/icp_2016-pt.pdf. Acesso em: 18 jul. 2021.

FUSCO, E. *Aplicação dos FRBR na modelagem de catálogos bibliográficos digitais*. São Paulo: Cultura Acadêmica, 2011.

GONZÁLEZ DE GOMEZ, M. N.; MACHADO, R. *A ciência invisível: o papel dos relatórios e as questões de acesso à informação científica*. *Pesquisa Brasileira em Ciência da Informação e Biblioteconomia*, v. 3, n. 1, 2008.

GRAVINA, R. A. S. Catalogação de objetos de moda: a roupa como instrumento de pesquisa. In: SEMINÁRIO DE INICIAÇÃO CIENTÍFICA DA FESPSP, 3., 2001, São Paulo. *Anais...* São Paulo: FESPSP, 2011. Disponível em: http://www.fespsp.org.br/sic2012/papers/2011/BIB/catalogacao_de_objetos_de_moda.pdf. Acesso em: 18 jul. 2021.

GUINCHAT, C.; M, M. *Introdução geral às ciências e técnicas da informação e documentação*. Brasília: IBICT, 1994. Disponível em: <http://livroaberto.ibict.br/handle/1/1007>. Acesso em: 18 jul. 2021.

LA FONTAINE, H.; OTLET, P. *Sur la création d'un répertoire bibliographique universel*. La Vve Larquier, 1895.

MARC 21: formato condensado para dados bibliográficos. Marília: UNESP: CGB Publicações Técnicas, 2000.

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

MEY, E. S. A.; SILVEIRA, N. C. Considerações teóricas aligeiradas sobre a catalogação e sua aplicação. *InCID: Revista de Ciência da Informação e Documentação*, Ribeirão Preto, v. 1, n. 1, p. 125-137, 2010.

PICADO, K. *Jarra bico de ganso*. João Pessoa, 2018. (Coleção: arquivo pessoal).

RIBEIRO, G. M.; CHAGAS, R. de L.; PINTO, S. L. *O renascimento cultural a partir da imprensa: o livro e sua nova dimensão no contexto social do século XV*. *Akrópolis: Revista de Ciências Humanas da UNIPAR, Umuarama*, v. 15, n. ½, p. 29-36, jan./jun. 2007.

SANTOS, J. L. *O que é cultura*. 6. ed. São Paulo: Brasiliense, 1987.

SANTOS, M. J. V. da C. *A representação da informação em Arquivos: viabilidade de uso dos padrões utilizados na Biblioteconomia*. *Acervo: Revista do Arquivo Nacional, Rio de Janeiro*, v. 20, n. 1-2, p. 57-66, jan./dez. 2007.

SOUZA, D. H. F. *Entidades coletivas: cabeçalhos e representação descritiva*. Belém: Cejup, 2003.

SOBRE AS AUTORAS E ORGANIZADORAS

Amanda Freitas dos Santos

Especialista em Gestão de Acervos e Unidades de Informação pela Universidade Tiradentes (Unit). Assistente Técnico no Centro de Documentação da Fundação Joaquim Nabuco (Fundaj).

Contato: amanda.santos-tc@fundaj.gov.br

ORCID: <https://orcid.org/0009-0006-3738-0088>

Ana Lúcia Terra

Doutora em Ciência da Informação pela Universidade de Coimbra (UC), Portugal. Docente na Faculdade de Letras da Universidade de Coimbra (UC).

Contato: anaterra@fl.uc.pt

ORCID: <https://orcid.org/0000-0003-1292-2849>

Brisa Pozzi de Sousa

Doutora em Ciência da Informação pela Universidade Federal de Minas Gerais (UFMG). Docente na Universidade Federal do Estado do Rio de Janeiro (UNIRIO).

Contato: brisa.pozzi@unirio.br

ORCID: <https://orcid.org/0000-0002-0298-8759>

Caterina Groposo Pavão

Doutora em Comunicação e Informação pela Universidade Federal do Rio Grande do Sul (UFRGS). Docente na Universidade Federal do Rio Grande do Sul (UFRGS).

Contato: caterina@ufrgs.br

ORCID: <http://orcid.org/0000-0003-3712-7200>

Cleusa Pavan

Doutora em Educação em Ciências pela Universidade Federal do Rio Grande do Sul (UFRGS). Bibliotecária na Universidade Federal do Rio Grande do Sul (UFRGS).

Contato: cleusa.pavan@ufrgs.br

ORCID: <https://orcid.org/0000-0002-1474-6669>

Darlene Alves Bezerra

Doutora em Ciência da Informação pela Universidade Federal Fluminense (UFF). Realizou estágio pós-doutoral em Letras na Universidade do Estado do Rio de Janeiro (UERJ) e Fundação Biblioteca Nacional (FBN).

Contato: darlenealvesbezerra@gmail.com

ORCID: <https://orcid.org/0000-0001-9353-5917>

Edilene Toscano Galdino dos Santos

Doutora em História e Filosofia da Ciência pela Universidade de Évora (UÉ), Portugal. Docente na Universidade Federal da Paraíba (UFPB).

Contato: edilenetoscanogs@gmail.com

ORCID: <https://orcid.org/0000-0002-1305-3289>

Elisabete Gonçalves de Souza

Doutora em História e Filosofia da Educação pela Universidade Estadual de Campinas (Unicamp). Docente na Universidade Federal Fluminense (UFF).

Contato: elisabetegs@id.uff.br

ORCID <https://orcid.org/0000-0001-9707-6017>

Giovana Deliberali Maimone

Doutora em Ciência da Informação pela Universidade de São Paulo (USP). Docente na Universidade de São Paulo (USP).

Contato: gdmaimone@usp.br

ORCID: <https://orcid.org/0000-0002-4150-8084>

Karin Herculano Picado

Doutoranda em História e Filosofia da Ciência pela Universidade de Évora (UÉ), Portugal. Mestra em Ciência da Informação pela Universidade Federal da Paraíba (UFPB). Detém experiência na área de Artes, com ênfase em Artes Gráficas, atuando principalmente em: xilogravura, serigrafia e ilustração à mão livre, curadoria de exposições.

Contato: karinpicado@gmail.com

ORCID: <https://orcid.org/0000-0001-6454-3598>

Luciane Paula Vital

Doutora em Ciência da Informação pela Universidade Federal de Santa Catarina (UFSC). Docente na Universidade Federal de Santa Catarina (UFSC).

Contato: luciane.vital@ufsc.br

ORCID: <https://orcid.org/0000-0003-2526-227X>

Lais Pereira de Oliveira

Doutora em Ciência da Informação pela Universidade Estadual Paulista "Júlio de Mesquita Filho" (Unesp). Docente na Universidade Federal de Goiás (UFG).

Contato: laispereira2@ufg.br

ORCID: <https://orcid.org/0000-0001-9092-4204>

Márcia Ivo Braz

Doutora em Ciências da Linguagem pela Universidade Católica de Pernambuco (Unicap). Docente na Universidade Federal de Pernambuco (UFPE).

Contato: marcia.ibraz@ufpe.br

ORCID: <https://orcid.org/0000-0003-2075-6094>

Naira Christofolletti Silveira

Doutora em Ciência da Informação pela Universidade de São Paulo (USP). Docente na Universidade Federal do Estado do Rio de Janeiro (UNIRIO) e Coordenadora-Geral do Centro de Pesquisa e Editoração da Fundação Biblioteca Nacional (FBN).

Contato: naira.silveira@unirio.br

ORCID: <https://orcid.org/0000-0002-0490-0052>

Patricia Andreia Lira

Mestra em Ciência da Informação, Bacharela e Licenciada em História pela Universidade de São Paulo (USP). Coordenadora da Área de Acervos da Associação Cultural Ciccillo Matarazzo, no Museu da Imagem e do Som de São Paulo (MIS) e no Paço das Artes.

Contato: anlira@gmail.com

ORCID: <https://orcid.org/0000-0003-1615-5755>

Paula Regina Dal'Evedove

Doutora em Ciência da Informação pela Universidade Estadual Paulista "Júlio de Mesquita Filho" (Unesp), Doutora em Gestión de la Información pela Universidad de Murcia (UM), Espanha. Docente na Universidade Federal de São Carlos (UFSCar) e Colaboradora no Programa de Pós-Graduação em Ciência da Informação da Unesp.

Contato: dalevedove@ufscar.br

ORCID: <https://orcid.org/0000-0003-1521-2201>

Raimunda Fernanda dos Santos

Doutora em Ciência da Informação pela Universidade Federal da Paraíba (UFPB). Docente na Universidade Federal do Rio Grande do Norte (UFRN). Docente permanente externa do Programa de Pós-Graduação em Ciência da Informação da Universidade Estadual de Londrina (PPGCI/UEL).

Contato: raimunda.fernanda@ufrn.br

ORCID: <https://orcid.org/0000-0002-7750-3269>

Rosali Fernandez de Souza

Doutora em Ciência da Informação pelo Polytechnic of North London, conferido pelo Council for National Academic Awards (PNL/CNAA), Inglaterra. Docente no Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT). Pesquisadora Titular aposentada do Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT), Ministério da Ciência, Tecnologia e Inovação (MCTI).

Contato: rosalifs@gmail.com

ORCID: <https://orcid.org/0000-0002-0890-7999>

Zaira Regina Zafalon

Doutora em Ciência da Informação pela Universidade Estadual Paulista “Júlio de Mesquita Filho” (Unesp). Docente na Universidade Federal de São Carlos (UFSCar).

Contato: zaira@ufscar.br

ORCID: <https://orcid.org/0000-0002-4467-2138>

Esta coletânea reúne pesquisadoras brasileiras que assumem, com densidade reflexiva e solidez teórica, a relação entre a Representação Documental e o ambiente digital, destacando os impactos, as possibilidades e os desafios que a era da digitalidade impõe ao campo da informação. Em um mundo cada vez mais interconectado, esta obra expressa a intenção de pensar e refletir sobre as transformações nas práticas de organização e representação da informação e do conhecimento, propondo a integração entre métodos tradicionais e abordagens inovadoras. Inovar, no que se refere à representação documental, pressupõe explorar e assumir a digitalidade e as novas demandas que ela traz. Voltada tanto para a comunidade acadêmica quanto para profissionais da área, esta obra oferece uma visão abrangente e atual das mudanças que estão redefinindo a Representação Documental na contemporaneidade.