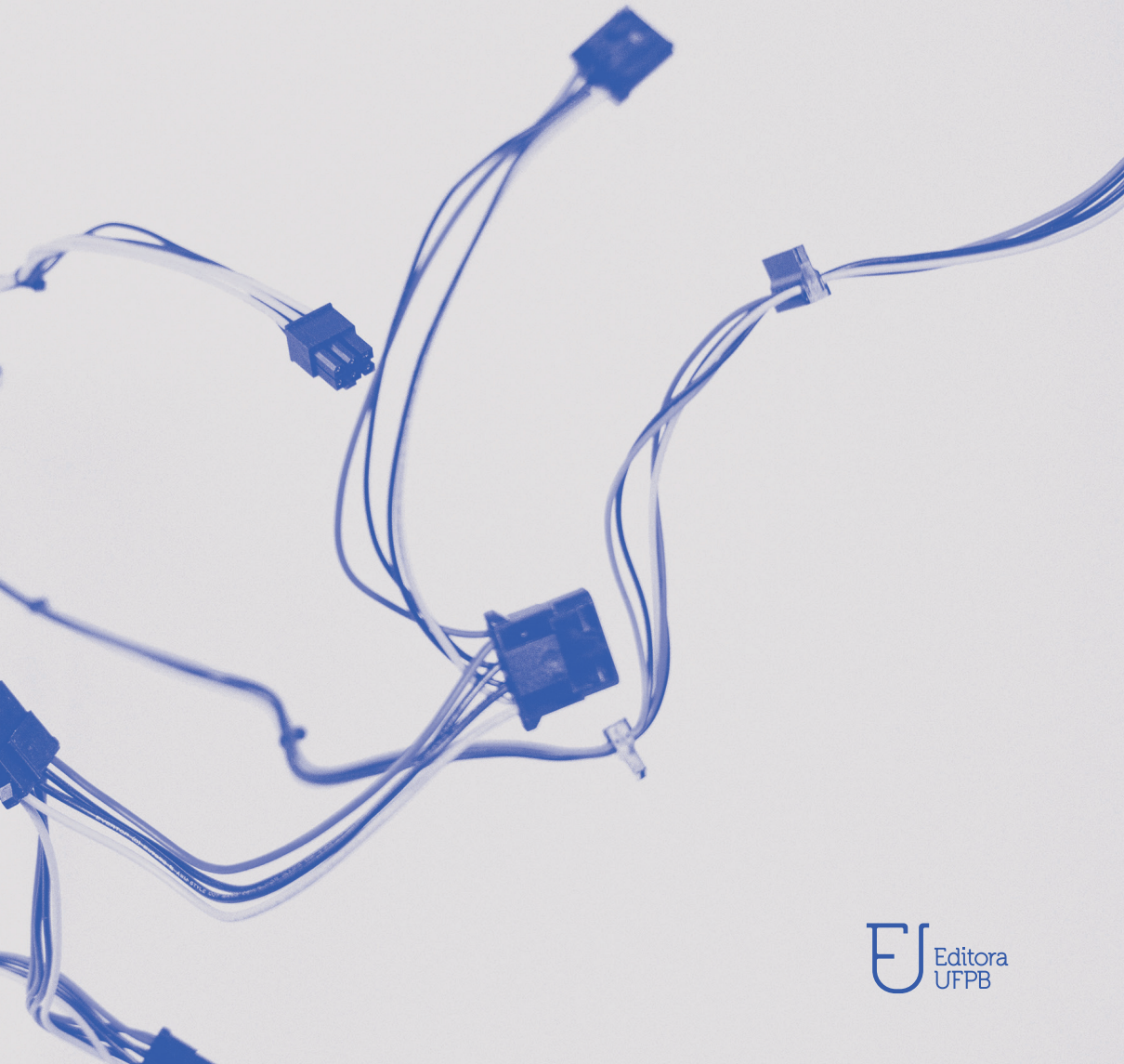
Henry Poncio Cruz de Oliveira
Silvana Aparecida Borsetti Gregorio Vidotti
(Organizadores)

INFORMAÇÃO E TECNOLOGIAS

DESENHANDO FRONTEIRAS CIENTÍFICAS



Editora UFPB

Informação e tecnologias

Desenhando fronteiras científicas

UNIVERSIDADE FEDERAL DA PARAÍBA

Reitora MARGARETH DE FÁTIMA FORMIGA MELO DINIZ

Vice-Reitora BERNARDINA MARIA JUVENAL FREIRE DE OLIVEIRA

Pró-Reitora PRPG MARIA LUIZA PEREIRA DE ALENCAR MAYER FEITOSA

EDITORA UFPB

Diretora IZABEL FRANÇA DE LIMA

Supervisão de Administração GEISA FABIANE FERREIRA CAVALCANTE

Supervisão de Editoração ALMIR CORREIA DE VASCONCELLOS JÚNIOR

Supervisão de Produção JOSÉ AUGUSTO DOS SANTOS FILHO

CONSELHO EDITORIAL

Adailson Pereira de Souza (Ciências Agrárias)

Eliana Vasconcelos da Silva Esvael (Linguística, Letras e Artes)

Fabiana Sena da Silva (Interdisciplinar)

Gisele Rocha Côrtes (Ciências Sociais Aplicadas)

Ilda Antonieta Salata Toscano (Ciências Exatas e da Terra)

Luana Rodrigues de Almeida (Ciências da Saúde)

Maria de Lourdes Barreto Gomes (Engenharias)

Maria Patrícia Lopes Goldfarb (Ciências Humanas)

Maria Regina Vasconcelos Barbosa (Ciências Biológicas)

Henry Poncio Cruz de Oliveira

Silvana Aparecida Borsetti Gregorio Vidotti

(Organizadores)

Informação e tecnologias

Desenhando fronteiras científicas

Editora UFPB

João Pessoa

2018

Direitos autorais 2018 – Editora UFPB

Efetuado o Depósito Legal na Biblioteca Nacional, conforme a Lei nº 10.994, de 14 de dezembro de 2004.

Todos os direitos reservados à Editora UFPB

É proibida a reprodução total ou parcial, de qualquer forma ou por qualquer meio.
A violação dos direitos autorais (Lei nº 9.610/1998) é crime estabelecido no artigo 184 do Código Penal.

O conteúdo desta publicação é de inteira responsabilidade dos autores.

Impresso no Brasil. Printed in Brazil.

Projeto Gráfico EDITORA UFPB

Editoração Eletrônica e Projeto da Capa ANA GABRIELLA CARVALHO

Imagem da Capa RANDALL BRUDER

Catalogação na fonte:
Biblioteca Central da Universidade Federal da Paraíba

I43 Informação e tecnologias : desenhando fronteiras científicas / Henry Poncio Cruz de Oliveira, Silvana Aparecida Borsetti Gregorio Vidotti (organizadores). - João Pessoa : Editora UFPB, 2018.

206 p.

ISBN 978-85-237-1328-7

1. Ciência da informação. 2. Repositórios digitais. 3. Informação e tecnologia. 4. Sistemas de informação. 5. Arquitetura da informação. I. Oliveira, Henry Poncio Cruz de. II. Vidotti, Silvana Aparecida Borsetti Gregorio. III. Título.

UFPB/BC CDU 02

EDITORA UFPB Cidade Universitária, Campus I, Prédio da Editora Universitária, s/n

João Pessoa – PB

CEP 58.051-970

http://www.editora.ufpb.br

E-mail: editora@ufpb.br

Fone: (83) 3216.7147

Editora filiada à: Associação Brasileira das Editoras Universitárias

Livro aprovado para publicação através do Edital Nº 4/2017-2018, financiado pelo Programa de Apoio a Produção Científica - Pró-Publicação de Livros da Pró-Reitoria de Pós-Graduação da Universidade Federal da Paraíba.

Sumário

Apresentação 07

Arquitetura da informação pervasiva:
modelo para repositórios digitais institucionais 09
Henry Poncio Cruz de Oliveira
Mayane Paulino de Brito e Silva

Integração entre serviços de repositórios digitais e
sistemas de informação de pesquisa corrente 25
Silvana Aparecida Borsetti Gregorio Vidotti
Carlos Alberto Baptista Sousa Pinto
Plácida Leopoldina Ventura Amorim da Costa Santos
Sandra Milena Roa Martínez
Felipe Augusto Arakaki

Ecologia informacional complexa das bibliotecas universitárias 49
Edgar Bisset Alvarez
Silvana Aparecida Borsetti Gregorio Vidotti
Juan Antonio Pastor-Sánchez

Perspectivas de investigação em encontrabilidade da informação:
em foco o atributo Descoberta de Informação 79
Fernando Luiz Vechiato

Análise, tendências e perspectivas da produção científica
em curadoria digital 99
Sandra de Albuquerque Siebra
Vildeane da Rocha Borba

Gestão da qualidade de dados conectados: fundamentos e metodologia .. 115
Leonardo Castro Botega
José Eduardo Santarem Segundo
Jéssica Oliveira de Souza Ferreira Melo
Caio Saraiva Coneglian

Ansiedade da informação revisitada: reflexões teóricas com base na psicologia e na ciência da informação 137
Henry Poncio Cruz de Oliveira
Josevânia da Silva

El lenguaje de la documentación en salud: algunos obstáculos evitables .. 149
Rosa Estopà Bagot

O Facebook como espaço tecnológico e discursivo d@s usuári@s 165
Daniel Thiago Freire da Silva
Izabel França de Lima
Maria da Luz Olegário

Verbete, entidade e *query*: em busca de um construto do conhecimento e da significação no ciberespaço 179
Silvana Drumond Monteiro
Gian Carlo Decarli

Universidades que participam desta obra 197

Autorxs 199

Apresentação

Uma fronteira pode ser compreendida como uma linha imaginária que cria uma divisão ou uma delimitação. Na medida em que separa, por meio de fios abstratos, uma fronteira faz uma demarcação que nos permite fazer a distinção entre um espaço e outro, entre um objeto e outro, entre uma coisa e outra (DICIO, 2018[1]). Nos contextos científicos, as fronteiras delimitam espaços entre o conhecido e o desconhecido, criam territórios de pesquisa e investigação, separam e ao mesmo tempo reúnem pesquisadores e proporcionam o crescimento de pesquisas sobre temáticas específicas.

A Ciência, num sentido mais clássico e amplo, pode ser compreendida como um conjunto articulado de conhecimentos "[...] metodicamente adquiridos, mais ou menos sistematicamente organizados, e suscetíveis de serem transmitidos por um processo pedagógico de ensino" (JAPIASSU; MARCONDES, 2008, p.44[2]). Estes conjuntos de conhecimentos científicos também são demarcados por linhas abstratas e constitutivas que delineiam os diversos campos e temáticas científicas.

Nesta obra, refletimos sobre algumas fronteiras científicas relacionadas ao articulado temático da "Informação e Tecnologia", no contexto da Ciência da Informação. Nosso ponto de partida foi a percepção de que as fronteiras científicas são dinâmicas e podem se modificar no processo de acumulação científica. Também nos embasamos no caráter interdisciplinar da Ciência da Informação, que, em nossa concepção, propulsiona novos olhares sobre o conhecido e uma gama de possíveis olhares sobre o desconhecido.

Neste contexto, reunimos pesquisas cujas temáticas nos colocam na fronteira da produção do conhecimento em "Informação e Tecnologia". Por meio desta obra, os leitores podem visitar aplicações e tendências em *Arquitetura da informação pervasiva*, compreender e visualizar as *Ecologias informacionais complexas*, avançar no entendimento da *Encontrabilidade da informação*, visitar

1 FRONTEIRA. In: DICIO, Dicionário Online de Português. Disponível em: <https://www.dicio.com.br/fronteira/>. Acesso em: 8 mar. 2018.

2 JAPIASSU, Hilton; MARCONDES, Danilo. Dicionário Básico de Filosofia. 5. ed. Rio de Janeiro: Jorge Zahar, 2008.

possiblidades de *Integração entre serviços de repositórios digitais e sistemas de informação de pesquisa corrente*.

As fronteiras da *Gestão da qualidade de dados conectados* e da *Curadoria digital* também podem ser visualizadas nesta obra. Considerando que uma fronteira se modifica ao longo do tempo e que na Ciência é possível estabelecer novos olhares científicos sobre temáticas, por vezes, já consolidadas, temos neste livro reflexões sobre *Ansiedade da informação*, sobre *Informação e linguagem em ambientes de saúde*, sobre *O Facebook como espaço tecnológico e discursivo* e sobre *Construto do conhecimento e de significação no ciberespaço*.

Reunimos pesquisadores de treze Universidades, sendo quatro estrangeiras, fortalecendo uma malha interinstitucional de pesquisas e pesquisadores comprometidos com o debate científico sobre "Informação e Tecnologia".

Convidamos você, caro leitor, a navegar nos mares deste livro, que cada capítulo traga uma experiência proveitosa e que a leitura possibilite o mergulho nas mais diversas fronteiras científicas informacionais e tecnológicas.

Henry Poncio Cruz de Oliveira

Silvana Aparecida Borsetti Gregorio Vidotti

Arquitetura da informação pervasiva: modelo para repositórios digitais institucionais

Henry Poncio Cruz de Oliveira
Mayane Paulino de Brito e Silva

1 Introdução

Diante das novas demandas da nossa sociedade, reflete-se neste trabalho sobre o que Resmini e Rosati (2011a) e Oliveira (2014) chamam de Arquitetura da Informação Pervasiva (AIP). Trata-se de um tema significativo, já que emerge um movimento tecnológico de integração de ambientes analógicos com ambientes de informação digital, por intermédio da criação de camadas informacionais que os entrecruzam, tornando a experiência dos usuários holística.

Oliveira (2014) assevera que a Arquitetura da Informação Pervasiva é uma temática fronteiriça e emergente, por isso necessita de aprofundamentos teóricos e metodológicos.

Para esta pesquisa, sob a perspectiva da AIP, buscou-se foco nos Repositórios Institucionais. Neste capítulo, esses repositórios foram referenciados como Repositórios Digitais Institucionais (RDIs).

Sendo assim, o objetivo deste estudo foi apresentar um modelo metodológico para avaliação da Arquitetura da Informação Pervasiva em Repositórios Digitais Institucionais.

2 Arquitetura da informação pervasiva

O termo Arquitetura da Informação (AI) foi fortemente divulgado pelo arquiteto americano Richard Saul Wurman, na década de 1970, ao comparar os problemas de arquitetura dos edifícios aos de armazenamento, organização e apresentação da informação. Assim, o arquiteto americano fez da AI seu objeto

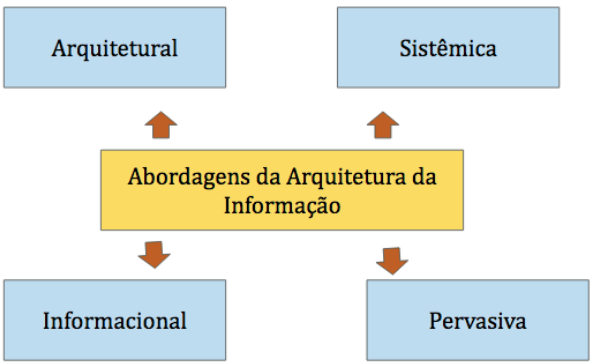
de estudo, objetivando organizar a informação de forma que seus usuários conseguissem acessá-la facilmente.

Wurman (1996) definiu a Arquitetura da Informação como a "ciência e a arte de criar instruções para espaços organizados".

A AI teria o propósito, então, de facilitar a vida dos usuários, permitindo a construção de espaços com boa usabilidade e acesso simples aos conteúdos.

Oliveira (2014), por meio de releituras de obras como León (2008) e Resmini e Rosati (2011a), chama a atenção para a influência de campos disciplinares na história da AI como a Arquitetura, Design, Ciência da Informação, Sistemas de Informação, Computação Ubíqua e Design de Serviços, os quais, por meio de seus próprios campos e seus núcleos conceituais, constituem as abordagens da AI, de acordo com o que é mostrado na Figura 1.

Figura 1 - Abordagens da Arquitetura da Informação



Fonte: Adaptado de Oliveira (2014).

Dentre estas abordagens, destaca-se, nesta pesquisa, a "pervasiva". Isto porque, atualmente, devido à nossa sociedade estar sendo direcionada para uma realidade de convergência entre os diversos ambientes de informação, os princípios da AI estão movimentando-se para zonas inexploradas, atingindo novas fronteiras e sendo conduzidos para além dos limites da *World Wide Web* (OLIVEIRA, 2014).

A abordagem pervasiva surge em razão do novo momento da Arquitetura da Informação, no qual é preciso investigar e apresentar soluções no âmbito tecnológico marcado pela pervasividade e ubiquidade, surgindo, portanto, o que Resmini e Rosati (2011a) nomearam de Arquitetura da Informação Pervasiva (AIP).

Oliveira (2014) descreve que os estudos de pesquisas em AIP avaliam, dentre outros aspectos, os processos de hibridização dos ambientes físicos utilizados pelos usuários.

Sousa e Pádua (2014, p. 70) salientam que:

> Diversos campos do conhecimento, práticas profissionais e disciplinas estão convergindo para a região de fronteira onde o design digital, recursos de rede, interações sociais e acesso aos dispositivos móveis se combinam, conduzindo a um questionamento sobre o foco em uma abordagem orientada pela informação, sendo justamente neste ponto que se iniciam as considerações sobre uma AIP.

Resmini e Rosati (2011b) acrescentam que, numa conjuntura em que ocorrem mudanças tecnológicas a todo instante, o comportamento dos usuários, do mesmo modo, modifica-se permanentemente. Nesse contexto, surge o termo "*prosumer*", o qual referencia os usuários que são, ao mesmo tempo, produtores e consumidores da informação. Considera-se o uso das tecnologias móveis como um catalisador desse processo, exigindo um redesenho das questões ligadas à organização, à representação, ao acesso, ao uso e à produção de informações digitais.

Resmini e Rosati (2011a) argumentam que nossas atividades diárias estão se modificando, de forma que nós nos movimentamos não somente de um meio para outro ou de um dispositivo para outro, mas em todas as esferas e direções. Um processo que se inicia digital acaba transformando-se em físico, de igual modo, um que começou físico torna-se digital.

Esses contextos refletem experiências *cross-channel*, as quais provocam reflexões em relação a AI em ambientes analógicos, digitais e híbridos, suscitando investigações para agregar físico e digital por meio do compartilhamento de

informações semelhantes, com o objetivo de amparar experiências informacionais holísticas, ecológicas e pervasivas, em conformidade com Oliveira (2014).

Contornando essas discussões, exibe-se o Quadro 1, o qual registra os fatos que contribuíram para o estudo da Arquitetura da Informação Pervasiva.

Quadro 1 - Importância para se estudar a Arquitetura da Informação Pervasiva

(continua)

As arquiteturas de informação transformaram-se em ecologias	Quando diferentes mídias e contextos relacionam-se fortemente, nenhum elemento pode ser encarado como uma entidade isolada, de modo que cada elemento passa a pertencer a um ecossistema maior. Por causa disso, deve ser considerado que todos esses elementos possuem vários links ou relacionamentos entre si, além de serem compreendidos como integrantes de um único processo no que diz respeito à experiência do usuário.
Os usuários são agora intermediários	Seguindo a ideia de "prosumers", percebe-se que os usuários tornaram-se participantes desses ecossistemas de informações, produzindo ativamente variados conteúdos nesses espaços.
A estática torna-se dinâmica	Isso quer dizer que as arquiteturas agregam conteúdos que, apesar de estarem presentes em ambientes físicos, podem ser libertados para finalidades de natureza diferente. Ademais, o papel de intermediários que os usuários passam a assumir faz com que os conteúdos se tornem perpetuamente inacabados, exigindo sempre novos refinamentos e manipulações.
O dinâmico torna-se híbrido	As novas arquiteturas englobam domínios distintos, do tipo físico, digital e híbrido, e de diferentes mídias. As experiências dos usuários passam a abranger diferentes ambientes.
O horizontal prevalece sobre o vertical	As tradicionais hierarquias de cima para baixo passam a dar lugar à dinâmica das novas arquiteturas em que a correlação entre os elementos passa a ser uma característica predominante.

Quadro 1 - Importância para se estudar a Arquitetura da Informação Pervasiva

(conclusão)

O design de produto passa a ser design de experiência	Levando em consideração que cada elemento (conteúdo, produto ou serviço) faz parte de um ecossistema macro, projetar experiências passa a ser mais ideal do que projetar itens únicos e isolados.
As experiências tornam-se experiências cross-media	Múltiplas experiências em ponte, além de variados meios de comunicação e ambientes pertencentes a ecologias onipresentes, contribuem para uma experiência de usuário sem uma delimitação global.

Fonte: Adaptado de Resmini e Rosati (2011a).

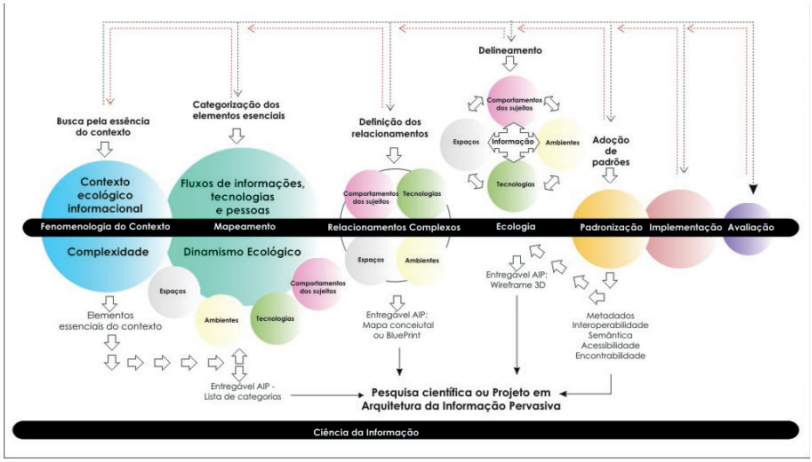
Concatenando as ideias apresentadas mediante um tratamento fenomenológico-sistêmico, Oliveira (2014, p. 166) conceitua a Arquitetura da Informação Pervasiva como:

> [...] uma abordagem teórico-prática da disciplina científica pós-moderna Arquitetura da Informação, trata da pesquisa científica e do projeto de ecologias informacionais complexas. Busca manter o senso de localização do usuário na ecologia e o uso de espaços, ambientes e tecnologias de forma convergente e consistente. Promove a adaptação da ecologia a usuários e aos novos contextos, sugerindo conexões no interior da ecologia e com outras ecologias. Facilita a interação com conjuntos de dados e informações ao considerar os padrões interoperáveis, a acessibilidade, a usabilidade, as qualidades semânticas e a encontrabilidade da informação, portanto deve buscar bases na Ciência da Informação.

Compreendendo os fatores que definem a AIP e objetivando materializar possibilidades para as investigações decorrentes de seu entendimento, Oliveira (2014) propôs um delineamento metodológico generalista para servir de base para a investigação da AI Pervasiva, tal como para nortear os arquitetos da

informação no que diz respeito a projetar, acompanhar e avaliar a Arquitetura da Informação Pervasiva em uma ecologia informacional complexa (Figura 2).

Figura 2 - Modelo metodológico para Arquitetura da Informação Pervasiva



Fonte: Oliveira (2014).

Sabendo que esse modelo é generalista e partindo para uma perspectiva aplicada da Arquitetura da Informação Pervasiva, nesta pesquisa é mostrado um modelo adaptado para o contexto dos Repositórios Digitais Institucionais, os quais são compreendidos pelos autores como ecologias informacionais complexas.

3 Metodologia

Com o interesse em seguir um modelo dinâmico, flexível e não reducionista para a resolução dos objetivos delineados nesta pesquisa, utilizou-se como fundamento o método quadripolar, inicialmente elaborado por Bruyne, Herman e Schoutheete, em 1991, com o intuito de construir um instrumento de pesquisa que se dirigisse para um novo paradigma nas Ciências Humanas e Sociais.

A opção pelo método quadripolar baseia-se nos argumentos de Silva e Ribeiro (2002), que propõem o método como um alicerce para investigação global na Ciência da Informação. Por sua natureza completa e dinâmica, ele torna-se

pertinente para ser utilizado nesta pesquisa, que é marcada pela complexidade e multidimensionalidade que envolvem a Arquitetura da Informação Pervasiva.

A dinâmica investigativa do método quadripolar resulta em quatro polos que interagem entre si: o polo epistemológico, o polo teórico, o polo técnico e o polo morfológico:

> No **pólo epistemológico**, opera-se a permanente construção do objecto científico e a definição dos limites da problemática de investigação, dando-se uma constante reformulação dos parâmetros discursivos, dos paradigmas e dos critérios de cientificidade que orientam todo o processo de investigação; no **pólo teórico**, centra-se a racionalidade do sujeito que conhece e aborda o objeto, bem como a postulação de leis, a formulação de hipóteses, teorias e conceitos operatórios e consequente confirmação ou informação do <contexto teórico> elaborado; no **pólo técnico**, consuma-se, por via instrumental, o contacto com a realidade objectivada, aferindo-se a capacidade de validação do dispositivo metodológico, sendo aqui que se desenvolvem operações cruciais como a observação de casos e de variáveis e a avaliação retrospectiva e prospectiva, sempre tendo em vista a confirmação ou refutação das leis postuladas, das teorias elaboradas e dos conceitos operatórios formulados; **no pólo morfológico**, formalizam-se os resultados da investigação levada a cabo, através da representação do objecto em estudo e da exposição de todo o processo de pesquisa e análise que permitiu a construção científica em torno dele (SILVA, 2006, p. 154- 155, grifo nosso).

Um estudo via método quadripolar implica a articulação e a interação entre os quatro polos, os quais não são peças estanques, isoladas ou separadas entre si, mas sim trabalham como um todo que se movimenta através de quatro engrenagens (SILVA, 2006).

No polo epistemológico, nesta pesquisa, há a demarcação da Arquitetura da Informação Pervasiva no cenário dos Repositórios Digitais Institucionais, problematizando-a como um fenômeno investigado pela Ciência da Informação no interior dialógico dos paradigmas cognitivo e social discutidos por Capurro (2003). Neste panorama, emergem os Repositórios Digitais Institucionais como objetos de investigação desta pesquisa, que ocorre no interior poliepistemológico na Ciência da Informação.

O polo teórico deste trabalho articula conceitos sobre Arquitetura da Informação Pervasiva e Repositórios Digitais Institucionais, a partir de referenciais da Ciência da Informação.

No polo técnico, temos como realidade objetivada que diz respeito aos RDIs. Tecnicamente, operou-se de modo a conferir a revisão de bibliografia e a adaptação do dispositivo metodológico de Oliveira (2014) para o contexto dos RDIs, numa perspectiva analítica, iterativa e sistêmica.

No polo morfológico, exibimos os quadros de análise por meio da capacidade sintética, flexível e dinâmica do método quadripolar. Neste sentido, destacamos que no polo morfológico atingimos o objetivo desta pesquisa ao apresentar o aparato metodológico de Oliveira (2014) adaptado para o contexto dos RDIs.

4 Repositórios digitais institucionais como ecologias informacionais complexas

Oliveira e Vidotti (2016) condensam a compreensão de espaço de informação como aquilo que tem o potencial de realizar a armazenagem dos objetos informacionais. Nesta conjuntura, as instituições de memória sempre se preocuparam com a seleção dos objetos que vão compor a sua coleção, no caso das bibliotecas, dos arquivos e dos museus, o interesse foi e é de criar coleções que possam representar o conhecimento acumulado (DODEBEI, 2009).

Hodiernamente, as universidades e os centros de pesquisa têm planejado e implantado repositórios digitais institucionais. Isso ocorre porque as comunidades acadêmicas estão participando de um movimento coletivo que busca explorar as

possibilidades oferecidas pelas Tecnologias Digitais de Informação e Comunicação (TDIC) para a criação de serviços de informação inovadores, por meio dos quais as novas mídias digitais possibilitam otimizar o ensino, a pesquisa e a comunicação científica (SAYÃO; MARCONDES, 2009).

Sabendo disso, uma das principais responsabilidades das instituições de conhecimento é promover a custódia sobre seus conteúdos produzidos, permitindo que eles estejam disponíveis para o seu acesso e para a sua preservação a longo prazo, segundo Sayão e Marcondes (2009).

Isto posto, os RDI responsabilizam-se por essa atribuição ao encarregarem-se da tarefa de assegurar que o conhecimento produzido por pesquisadores da instituição seja armazenado, disseminado e preservado.

Os repositórios digitais institucionais retiram a exclusividade das editoras de periódicos e passam para a instituição a tarefa de promover a guarda e o acesso a parte da massa de informação produzida no contexto das instituições.

Da perspectiva conceitual, pode-se compreender um repositório digital institucional como:

> [...] uma base de dados na Web na qual uma instituição de pesquisa deposita sistematicamente sua produção acadêmica e a disponibiliza de forma ampla para as comunidades interessadas. Sobre essa base de dados é oferecido um conjunto de serviços voltados para a gestão e para a disseminação de informações em formato digital. Esses serviços incluem captura, armazenamento, tratamento técnico, organização, preservação e entrega de conteúdos digitais de toda a natureza – texto, imagens, vídeo, áudio, apresentações, programas de computador, *datasets etc* (SAYÃO; MARCONDES, 2009, p. 23).

Os autores supracitados ainda observam que, como o próprio nome faz lembrar, os repositórios digitais institucionais estão ligados naturalmente aos seus ambientes institucionais, tendo um evidente compromisso com a formação da memória digital acadêmica e com a preservação de materiais a longo prazo.

Para isso, esses repositórios utilizam um suporte tecnológico de sistemas abertos e interoperáveis, bem como a aplicação de padrões das áreas de Biblioteconomia, Ciência da Informação e Tecnologia da Informação.

Com essas ponderações, entende-se o RDI como uma ecologia informacional complexa. Pautando-se em Oliveira (2014), o termo "ecologia" ganha ressignificação na Arquitetura da Informação Pervasiva ao delimitar as variadas relações entrecruzadas de pessoas, de processos e de elementos dos ambientes informacionais, com a finalidade de administrar a complexidade e a variedade do uso atual da informação. Por causa disso, no contexto da AIP, há um esforço para tratar objetos e fenômenos como uma estrutura informacional ecológica, sistêmica e complexa.

A estrutura informacional ecológica que é tratada nesta pesquisa refere-se aos espaços, ambientes, canais, mídias, tecnologias e usuários com seus comportamentos que estão interligados ao contexto dos Repositórios Institucionais.

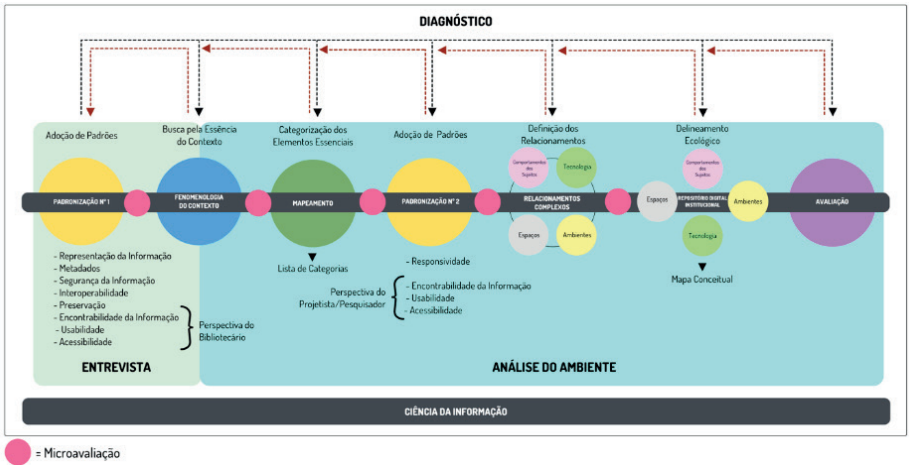
Pode-se visualizar essa ecologia abarcando, por exemplo, a página web de um Repositório Digital Institucional, as políticas relacionadas ao RDI, o site da instituição à qual esse repositório está ligado, suas redes sociais (Facebook, Instagram, Twitter *etc*), o catálogo da biblioteca da instituição, o site desta biblioteca, o ambiente físico desta biblioteca, os dispositivos utilizados para acessar os ambientes informacionais (*desktop*, *notebook*, *tablet*, *smartphone etc*), as tecnologias necessárias para permitir o acesso aos ambientes, além, claro, de todos os usuários, e suas respectivas necessidades e comportamentos, envolvidos no contexto do RDI (bibliotecários, estudantes, servidores, professores, comunidade vinculada à instituição em geral *etc*).

Nessa ecologia persiste o sistemismo de ambientes e relações complexas que ocorrem de forma intra e extraecológicas (OLIVEIRA, 2014).

5 Refletindo sobre a arquitetura da informação pervasiva em repositórios digitais institucionais

Tomando como subsídio o modelo metodológico proposto por Oliveira (2014), adaptou-se a organização do modelo para a realidade dos repositórios. Na adaptação realizada (Figura 3), uma das principais necessidades percebidas para o estudo da AIP nos RDI foi a utilização de uma entrevista semiestruturada para a compreensão de alguns aspectos relacionados ao objeto de estudo que não poderiam ser compreendidos com apenas a observação do pesquisador/projetista, uma vez que dependeria da visão do profissional responsável pelo gerenciamento do RDI.

Figura 3 - Modelo metodológico adaptado para a avaliação da AIP nos Repositórios Digitais Institucionais



Fonte: Modificado de Oliveira (2014).

O processo adaptativo mostrado na Figura 3 se baseou na compreensão do funcionamento básico de um RDI a partir da literatura científica da área de Ciência da Informação sobre repositórios.

No modelo adaptado, há dois momentos ou blocos metodológicos que merecem destaque: a entrevista e a análise do ambiente.

A entrevista tem o potencial de fornecer as características peculiares/contextuais do RDI e dos demais ambientes, características essas não possíveis de serem encontradas apenas por meio da observação do sistema. Essa entrevista não está prevista no modelo de Oliveira (2014), porém foi constatado que, principalmente no que tange às informações sobre a interoperabilidade do RDI, é essencial o diálogo com o profissional que trabalha diretamente com os ambientes elencados para análise.

No ponto referente à entrevista, há as fases chamadas de "padronização nº 1" e de "fenomenologia do contexto", sendo a última também imersa no segundo momento, em razão da necessidade tanto da visão do bibliotecário do RDI quanto da visão do pesquisador, quando realizada a análise do ambiente. A fenomenologia do contexto interseciona os blocos metodológicos da entrevista e da análise do ambiente.

A "padronização nº 1" investiga os padrões que são utilizados, sendo neste modelo adaptado apresentadas questões relacionadas, sobretudo, à interoperabilidade e ao padrão de metadados do RDI, além da investigação sobre representação da informação, segurança da informação, preservação, responsividade, usabilidade, encontrabilidade da informação e acessibilidade. Esta etapa objetivou estabelecer os recursos que permitem o trânsito de informações representáveis, seguras, preservadas, responsivas, usáveis, encontráveis e acessíveis às pessoas que interagem na ecologia. Destaca-se que ela foi inserida para ser o começo da investigação porque somente através do conhecimento de todos esses recursos por meio da entrevista foi possível enxergar quais são alguns dos elementos essenciais para o desempenho do RDI, o que leva à segunda etapa.

Na "fenomenologia do contexto" foram investigados, no contexto empírico, os ambientes e espaços relacionados ao RDI, as possibilidades tecnológicas aplicáveis e os usuários que utilizam o repositório. O termo fenomenologia, nesta etapa, é herdado das concepções e escolhas epistemológicas de Oliveira (2014), sendo empregado no sentido de se buscar os elementos essenciais e indispensáveis relacionados aos ambientes, espaços, tecnologias e usuários. A fenomenologia do contexto está inserida na entrevista, tendo em vista a busca da percepção do profissional que atua no repositório a respeito dos elementos

que, segundo a ideia dele, são indispensáveis para o funcionamento adequado do RDI. Ao mesmo tempo, a fenomenologia adentra o segundo momento, relativo à análise do ambiente, posto que é fundamental se ter a percepção real de quais são as essências do sistema de informação estudado.

Seguindo o fluxo do modelo, temos a fase do "mapeamento", que, baseando-se nas fases anteriores, torna possível mapear e categorizar espaços, ambientes, tecnologias e comportamentos ecológicos. Essa etapa visa explanar as alternativas de navegação do usuário em todos os sistemas que fazem parte do contexto holístico do RDI, por intermédio de uma lista de categorias.

Em seguida, vê-se a "padronização nº 2", que se diferencia da "padronização nº 1" pelo fato daquela levar em consideração a observação do pesquisador e do arquiteto da informação, em nível de projeto executado, quanto aos pontos relacionados à responsividade, à usabilidade, à encontrabilidade da informação e à acessibilidade. A padronização nº 2 não pondera, a priori, sobre interoperabilidade, metadados, representação da informação, segurança da informação e preservação, por ter sido refletido aqui que essas pontuações mais "abstratas" do sistema só podem ser respondidas profundamente pelo profissional que trabalha diretamente com ele, o que inclusive fez da entrevista uma parte essencial para se entender isso.

A etapa que constrói os "relacionamentos complexos" entre as partes da ecologia sugere as possíveis conexões intraecológicas e conexões com outras ecologias.

Nesse ponto, defronta-se com a epistemologia sistêmica do projeto da ecologia informacional complexa, pois as partes da ecologia tomam forma e sugerem a constituição de um todo estruturado por suas partes.

Essa etapa trouxe um esboço dos tipos de relacionamentos existentes na ecologia do RDI, bem como dos relacionamentos que, caso não sejam considerados na realidade, deveriam ser.

Feito isso, tem-se o estágio do "delineamento ecológico", no qual foi possível visualizar as diversas dimensões da ecologia, percebendo as camadas informacionais encarregadas de conectar as partes da ecologia, gerando um todo complexo e dinâmico, o qual foi compreendido através de um mapa conceitual que exibiu o resultado fruto dessa etapa.

Considerou-se a "avaliação" final para a aplicação do modelo adaptado para o contexto dos RDI. Ressalta-se que, ao longo de todo processo, como mostra a Figura 3, foram realizadas "microavalições" ao final de cada etapa como meio de avaliar individualmente o que foi alcançado em cada uma delas. Frisa-se que esse processo como um todo é qualificado como um diagnóstico de uma Arquitetura da Informação Pervasiva em Repositórios Digitais Institucionais.

Por fim, pontua-se que nenhuma etapa do modelo é estanque e a sequência exibida não é linear, sendo possível que ao término de uma etapa haja a necessidade de se voltar à anterior para uma conversa com os resultados obtidos, caracterizando o modelo adaptado para os RDI como iterativos. Por esses motivos que as microavaliações inseridas no modelo adaptado são tão importantes.

6 Considerações finais

No transcorrer do percurso desta pesquisa, relacionou-se a temática da Arquitetura da Informação Pervasiva com os Repositórios Digitais Institucionais, permitindo a inter/transdisciplinaridade por meio do diálogo entre a Ciência da Informação e a Arquitetura da Informação, tendo os RDI como ponto norteador.

Ao tentar responder a questão motivadora desta pesquisa, constatou-se que a Arquitetura da Informação Pervasiva pode ser visualizada e aplicada no contexto dos Repositórios Digitais Institucionais por intermédio do modelo adaptado da AIP para os RDIs.

Salienta-se que os Repositórios Digitais Institucionais possuem uma estrutura ecológica com uma diversidade de ambientes, espaços, tecnologias e usuários que contribuem para a aplicação da AIP, contudo não permitem a tendência de movimento, propagação, infiltração e difusão total desses elementos dentro da ecologia. Em virtude disso, não possibilitam que as informações que por eles circulam sejam trabalhadas holisticamente por meio de todos os componentes da ecologia informacional, levando em consideração as necessidades e subjetividades dos usuários. No entanto, mesmo que os RDI ainda não se disponham a essa realidade, vê-se que a pervasividade nos repositórios não é

algo distante, visto que eles já nasceram no contexto digital e, assim sendo, são facilmente adaptáveis às exigências atuais.

Referências

BRUYNE, P. de; HERMAN, J.; SCHOUTHEETE, M. de. **Dinâmica da pesquisa em ciências sociais**. 5. ed. Rio de Janeiro: Francisco Alves, 1991.

CAPURRO, R. **Epistemologia e ciência da informação**. 2003. Disponível em: http://www.capurro.de/enancib_p.htm. Acesso em: 11 set. 2017.

DODEBEI, V. Repositórios institucionais: por uma memória criativa no ciberespaço. *In*: SAYÃO, L. *et al*. (org.). **Implantação e gestão de repositórios institucionais**: políticas, memória, livre acesso e preservação. Salvador: EDUFBA, 2009.

LEÓN, R. R. Arquitectura de información: análisis histórico-conceptual. **No sólo usabilidade**, n. 7, abr. 2008. Disponível em: http://www.nosolousabilidad.com/articulos/historia_arquitectura_informacion.htm. Acesso em: 09 mai. 2019.

OLIVEIRA, H. P. C. de. **Arquitetura da informação pervasiva**: contribuições conceituais. 2014. Tese (Doutorado em Ciência da Informação) – Universidade Estadual Paulista "Júlio de Mesquita Filho", Marília, 2014.

OLIVEIRA, H. P. C. de.; VIDOTTI, S. A. B. G. Dos ambientes informacionais às ecologias informacionais complexas. **Inf. & Soc.: Est.**, João Pessoa, v. 26, n. 1, p. 91-101, jan./abr. 2016.

RESMINI, A.; ROSATI, L. A brief history of information architecture. **Journal of information architecture**, v. 3, n. 2, p. 33-45, 2011b.

RESMINI, A.; ROSATI, L. **Pervasive information architecture**: designing cross-channel user experiences. Burlington: Elsevier, 2011a.

SAYÃO, L. F.; MARCONDES, C. H. Introdução: repositórios institucionais e livre acesso *In*: SAYÃO, L. *et al*. (org.). **Implantação e gestão de repositórios institucionais**: políticas, memória, livre acesso e preservação. Salvador: EDUFBA, 2009.

SILVA, A. M. da. **A informação**: da compreensão do fenômeno e construção do objecto cientifico. Porto: Edições Afrontamento, 2006.

SILVA, M. A.; RIBEIRO, F. **Das "ciências" documentais à ciência da informação**: ensaio epistemológico para um novo modelo curricular. Porto: Afrontamento, 2002.

SOUSA, M. R. F. de; PÁDUA, M. C. Arquitetura da informação pervasiva: desvendando as heurísticas de Resmini e Rosati. **Informação & Tecnologia**, João Pessoa, v. 1, n. 1, p. 68-80, jan./jun., 2014.

WURMAN, R. S. **Information architects**. Zurich, Switzerland: Graphis Press, 1996.

Integração entre serviços de repositórios digitais e sistemas de informação de pesquisa corrente

Silvana Aparecida Borsetti Gregorio Vidotti
Carlos Alberto Baptista Sousa Pinto
Sandra Milena Roa Martínez
Felipe Augusto Arakaki
Plácida Leopoldina Ventura Amorim da Costa Santos

1 Introdução

A necessidade de determinados serviços impulsiona demandas de diferentes ramos de atuação. Nesse cenário, os serviços oferecidos em ambientes digitais têm aumentado proporcionalmente com a inserção de indivíduos em um mundo cada vez mais globalizado e conectado.

No âmbito das bibliotecas e centros de pesquisa, muitos serviços que antes eram feitos em suas localidades tradicionais agora foram transportados para o ambiente digital, principalmente no que tange a disponibilização e acesso a documentos e informações científicas. Assim, torna-se fundamental um sistema de gerenciamento de informação como repositórios digitais (RD) e *Current Research Information System* (CRIS), para organizar, preservar, recuperar, acessar e disseminar seus conteúdos para o público geral ou um público restrito.

A escolha do sistema de gerenciamento de informação deve levar em consideração diversos fatores, entre eles: a estrutura organizacional da instituição, os requisitos funcionais desejáveis do sistema, a arquitetura da informação, os esquemas de representação da informação, entre outros aspectos. Entretanto, mesmo com um planejamento bem definido, com o passar do tempo, a estrutura organizacional da instituição, suas prioridades, perspectivas e necessidades informacionais podem ser alteradas, assim como, as tecnologias podem evoluir, apresentando ferramentas mais adequadas para auxiliar determinados serviços.

Nesse contexto, é exigido dos gestores de sistemas de informação uma frequente atualização de seus conhecimentos, no que diz respeito ao entendimento sobre o comportamento e as necessidades informacionais de seu público alvo e novas ferramentas para ofertar serviços mais bem estruturados.

Dessa forma, com o propósito de esclarecer alguns pontos para gestores de sistemas de informação, questionou-se: qual a natureza dos serviços que podem ser oferecidos a partir da integração dos CRIS e repositórios digitais em relação às características informacionais e tecnológicas de uma instituição?

Assim, o objetivo deste artigo foi refletir, a partir da integração entre repositórios digitais e CRIS, sobre os serviços que podem ser suportados para gestão da informação de pesquisa com o intuito de analisar a natureza desses serviços.

Utilizou-se uma metodologia exploratória, uma vez que visou maior compreensão e familiaridade com as questões levantadas a partir de revisão de fundamentos teóricos, artigos científicos, documentos e projetos realizados anteriormente e relacionados com serviços em repositórios digitais institucionais e sistemas de informação para pesquisa. A técnica utilizada foi levantamento bibliográfico e revisão da literatura.

A estrutura deste documento contém os principais temas em que o trabalho proposto se baseou, tais como: Ciência de Serviços, Repositórios Digitais e CRIS. Em seguida, foram descritas as configurações da integração entre os repositórios digitais e os sistemas de informação para pesquisa, e apresentou-se uma análise da natureza dos serviços desta integração, ao final foram expostas as conclusões do trabalho realizado.

2 Ciência de serviços e serviços de informação

A composição de uma nova área é constituída por diversos fatores como: perspectivas de olhar sobre um objeto, técnicas de análise e metodologias de estudo. Essa demanda surge a partir da necessidade de entender fenômenos emergentes e principalmente com a interdisciplinaridade dos estudos, uma vez

que são necessários diversos conhecimentos para compreender aspectos e pontos de vista de um determinado objeto.

Nesse contexto, com a Ciência de Serviços não foi diferente. Conforme relatado no manifesto da Ciência de Serviços de Chesbrough e Spohrer (2006), a estruturação da Ciência de Serviços vem efetivando-se gradualmente, junto com as necessidades de entender as relações dos serviços com outros fatores. Spohrer e Kwan (2009) definem diversos componentes da Ciência de Serviços como recursos, entidades de sistemas de serviço, direitos de acesso, cocriação de valor, mecanismos de governança, redes de sistemas de serviços, ecologia de sistemas de serviços e resultados.

De acordo com Gianesi e Corrêa (1994) e Santos, Fachin e Varvakis (2003), os serviços estão ligados à vivência do usuário, portanto, são intangíveis e de difícil mensuração. Lovelock e Wright (2001) e Santos, Fachin e Varvakis (2003, p. 86) complementam que o serviço é caracterizado pelo "[...] ato ou desempenho que cria benefícios para clientes por meio de uma mudança desejada no – ou em nome do – destinatário do serviço." Segundo Maglio *et al.* (2009, p. 396), "Um comportamento-chave é que os sistemas de serviço interagem para co-criar valor." Spohrer e Maglio (2010, p. 3) definem de modo mais detalhado,

> [...] o serviço como co-criação de valores, o valor como mudança que as pessoas preferem e valorizam a co-criação como uma mudança ou conjunto de mudanças relacionadas que as pessoas preferem e realizam como resultado de sua comunicação, planejamento ou outro propósito e conhecimento intensivo de Interações.

Com o desenvolvimento e interdisciplinaridade das áreas e com o fenômeno emergente dos serviços, há cada vez mais domínios do conhecimento envolvidos, como: a inovação em serviços, gestão, marketing, engenharia, design, ciência da computação, ciência da informação, entre outros. A Ciência de Serviços, pode "[...] integrar-se em muitas disciplinas que já estudam aspectos do mundo

artificial, interagindo com as entidades de sistemas de serviços." (SPOHRER; MAGLIO, 2010, p. 6).

Assim, a Ciência de Serviços:

> [...] busca criar um corpo de conhecimento que represente a co-criação de valor entre as entidades à medida que elas interagem - descrever, explicar e (talvez um dia) prever, controlar e guiar melhor a evolução dos fenômenos de valorização (SPOHRER; MAGLIO, 2010, p. 3).

Relativamente nova, a Ciência de Serviço, vem sendo mais utilizada a partir do ano 2000 em simpósios científicos e eventos empresariais, destacando a importância da pesquisa e educação no campo de serviços (CHESBROUGH, 2004; CALABRIA *et al*., 2013). Spohrer e Maglio (2010) resumem que a Ciência de Serviços tem o objetivo de ajudar na criação das melhores e mais eficazes relações de cocriação de valor. Para Calabria *et al*. (2013, p. 118),

> Um dos objetivos principais dentro da Ciência de Serviços é o estudo da inovação em serviços e os mecanismos e modelos para a criação de novos serviços, seja de forma contínua, seja de forma descontínua ou ainda disruptiva.

Conforme Spohrer e Maglio (2008; 2010), a Ciência de Serviços tem como objetivo ainda explicar e melhorar as interações entre as várias entidades para alcançar resultados vantajosos para o benefício mútuo. A Ciência de Serviços está diretamente relacionada com a cocriação de valor, ou seja, é necessária a interação do cliente com um bem ou serviço para que assim gere um valor e consequentemente a cocriação. As especificidades dos estudos dos serviços são definidas pelas perspectivas em que os fenômenos ocorrem. Nesse contexto, entre os tipos de serviços estão os relacionados diretamente com aqueles oriundos das Tecnologias de Informação e Comunicação (TIC), denominados serviços de informação.

Outrossim, para Dholakia, Mundorf e Dholakia (1997), os serviços de informação possuem características únicas, além daquelas características relacionadas aos serviços. As características únicas dos serviços de informação são: serviços em rede, interatividade e caráter externo à rede. Os serviços em rede e a interatividade criam uma característica nos serviços de informação, que é o caráter externo da rede (KATZ; SHAPIRO, 1986; DHOLAKIA; MUNDORF; DHOLAKIA, 1997). Por exemplo, no contexto das bibliotecas, os serviços de informação sempre estiveram atrelados ao serviço de referência, à Disseminação Seletiva de Informação (DSI), ao acesso à Internet, à comutação bibliográfica, ao acesso a bases de dados internacionais, à resposta de questões, ao atendimento e consultas técnicas, entre outros serviços (BARROS, 2003).

3 Repositórios digitais

No contexto dos serviços, os repositórios digitais (RD) podem agregar diversos serviços de informação para ampliar a interação e participação do usuário, que serão abordados nesta seção.

De acordo com Santarem Segundo (2010, p. 151), os repositórios digitais são

> [...] sistemas de informação que facilitam a publicação e o armazenamento de documentos, além de fornecer serviços de informação, e por isso o interesse em contribuir com a organização de sua informação.

Entre as principais características dos repositórios digitais estão a interoperabilidade e o autoarquivamento. Basicamente, a interoperabilidade corresponde à troca de informações entre sistemas. Já o autoarquivamento é caracterizado pelo "depósito efetuado pelos próprios pesquisadores de suas respectivas produções científicas em repositórios digitais de acesso aberto." (FERRARI; PIRES, 2014, p. 24). Nessa perspectiva, os repositórios digitais podem ser institucionais ou temáticos.

Os repositórios institucionais caracterizam-se pela reunião da produção científica, artística, cultural e administrativa da instituição. Leite (2009, p. 21) complementa que

> [...] um repositório institucional consiste em um conjunto de serviços que a universidade oferece para os membros da sua comunidade com vista ao gerenciamento e disseminação do material digital criado pela instituição e pelos seus membros.

A reunião de documentos de uma área do conhecimento ou disciplina é denominada de repositório temático. Para Kuramoto (2006, p. 83), um repositório temático é:

> [...] um conjunto de serviços oferecidos por uma sociedade, associação ou organização, para gestão e disseminação da produção técnico-científica em meio digital, de uma área ou subárea específica do conhecimento.

A partir das características destacadas por Ferreira (2008), podem encontrar-se implícitos ou explícitos alguns dos serviços que um repositório oferece:

a) mecanismo de recuperação contextualizada;
b) padrões de organização, gerenciamento e publicação de conteúdos digitais;
c) preservação digital dos conteúdos;
d) sistema de gestão mediante integração com outros serviços;
d) regras, normas e padrões para armazenamento, preservação, divulgação e acesso da produção científica;
e) validação das autorias e qualificações correspondentes;
f) visibilidade da produção da instituição e/ou de uma área do conhecimento;
g) preservação dos direitos autorais em longo prazo;
h) materiais em distintas fases de publicação (preprints, post-prints e materiais publicados internamente);
i) materiais total ou parcialmente abertos, revisados por pares ou não;
j) inúmeros suportes (vídeos, filmes, textos, multimídia, fotos);

k) distintos formatos (como ppt, pdf, txt e jpeg, entre outros);
l) diversidade de tipos de documentos (artigos, livros, documentos de eventos, teses e dissertações, materiais didáticos, etc.).

Adicionalmente, diante da necessidade de disponibilização e visibilidade da produção científica, os repositórios digitais têm se tornado também uma alternativa para órgãos governamentais, instituições e centros de pesquisa reunir, em um ambiente, documentos de interesse da sua comunidade ou do público em geral.

Entretanto, com o decorrer dos anos, algumas iniciativas têm alterado as características iniciais das plataformas, como o caso do Repositório Institucional da Universidade Estadual Paulista (Unesp), que implementou um procedimento novo de reuso de dados ao invés do autoarquivamento. A Unesp desenvolveu uma metodologia para o povoamento automático de repositórios digitais,

> [...] que se deu por meio da inserção de registros em lote a partir da coleta automática dos metadados de publicações científicas (objetos digitais) em diferentes bases indexadoras e fontes de informação (plataformas de busca e portais) (VIDOTTI *et al*., 2016, p. 2).

A coleta automática poupa o tempo do pesquisador, que não vai precisar inserir as informações de suas publicações no repositório, reduz também o tempo do revisor que teria que verificar se todas as informações foram preenchidas corretamente. Ao serem migrados de outras bases de dados, os dados já foram descritos e revisados pelos gestores dessas bases, fato que dispensa nova revisão e inclusão no repositório.

Outra possibilidade que pode ser adotada pelos repositórios é a inclusão de serviços de atendimento ao usuário via *chat*, como alternativa ao e-mail ou ao telefone. Os serviços de disseminação seletiva da informação também poderão ser ampliados, combinando-os com os serviços de recomendações, analisando e comparando o perfil do usuário, seu comportamento no ambiente digital e mapeando suas necessidades informacionais.

Diversos outros serviços podem ser oferecidos em repositórios digitais, dependendo dos dados e das informações que uma instituição pode criar e gerenciar. Nessa perspectiva, as instituições necessitam, cada vez mais, aprimorar seus sistemas e encontrar alternativas para gestão da informação de sua unidade. Em alguns contextos, instituições precisam de sistemas que ajudem a gerenciar desde o início de uma pesquisa (um projeto) até seu produto final (como um artigo publicado em um periódico, uma tese ou uma dissertação). Nesse âmbito, surgiram os *Current Research Information System*, mais conhecidos como CRIS.

4 Current research information system

Os aspectos relacionados com as atividades do processo de pesquisa científica desenvolvidas em uma instituição geram dados que, uma vez inter-relacionados e analisados, podem fornecer informação útil na tomada de decisões associadas com: o recurso humano participante, os projetos, os resultados ou produtos e recursos financeiros e de investimento, entre outros. Esses dados são armazenados em *Current Research Information System*.

O CRIS pode ser entendido como um sistema de informação que coleta e armazena em um banco de dados os metadados, detalhando as entidades (como: projetos, pesquisadores, financiadores e publicações) e os relacionamentos participantes nos fluxos de trabalho próprios das atividades de pesquisa de uma instituição, bem como as interações que se estabelecem entre as entidades (SALES; SAYÃO, 2015).

Ampliando essa interpretação, a Organização Europeia para Informação Internacional de Pesquisa (euroCRIS) define um CRIS como um sistema que:

> [...] contém dados / metadados ou informações sobre gerentes de projetos, projetos em andamento e concluídos, departamentos de pesquisa, organizações de financiamento, programas e financiamentos, pesquisadores, resultados de pesquisas (publicações, patentes, produtos), eventos, instalações, (Semântica) e fornece uma abordagem integrada

> para gerenciar informações de pesquisa. (euroCRIS, 2010, não paginado, tradução nossa).

É importante reconhecer que a pesquisa científica em seu ciclo de vida, envolve vários elementos que tornam complexa a disseminação da informação, entre esses elementos encontram-se também, segundo Zimmerman (2002), pessoas, projetos, organizações, resultados de pesquisas (publicações, patentes e produtos), instalações e equipamentos, que são abrangidos na definição de CRIS. As ferramentas de software integram esses diversos elementos, e, a partir deles, oferecem uma gama de benefícios e/ou serviços que, de forma isolada, não poderiam ser concebidos isoladamente.

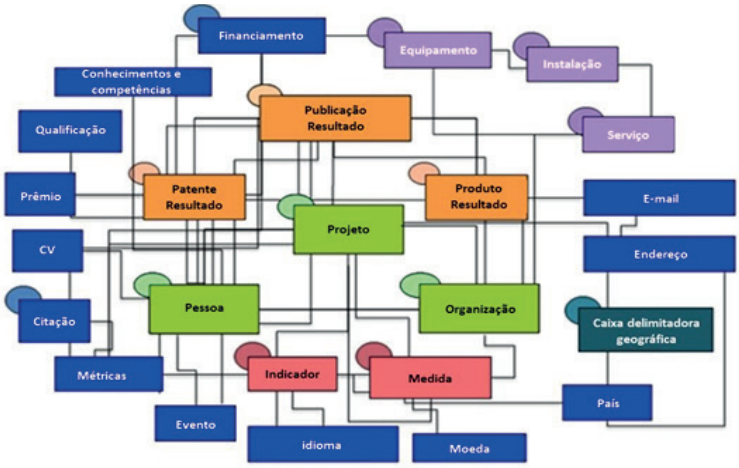
Além disso, os CRIS têm diferentes tipos de usuários, eles são: pesquisadores, gestores de pesquisa, agências de financiamento à pesquisa, formuladores de política de CT&I, empreendedores e inovadores, professores e estudantes, órgãos de divulgação e cidadão comum (SALES; SAYÃO, 2015).

Existem diferentes tipos de CRIS, caracterizados segundo o seu âmbito como: institucionais, nacionais, regionais e internacionais. Há também CRIS classificados por temática, ou seja, por área de conhecimento, objeto de estudo ou assunto (PINTO; SIMÕES; AMARAL, 2014).

O tipo de CRIS existente ou que busca desenvolver-se em uma instituição ou instituições contemplará as necessidades destes participantes, assim como os serviços que este sistema oferecerá estarão diretamente relacionados e dependentes do tipo de usuário.

Alguns CRIS são projetados e implementados como soluções próprias que atendem às necessidades específicas de uma instituição pelas peculiaridades que ela pode ter sobre as atividades do ciclo de vida da pesquisa. Além disso, também existem CRIS comerciais de apoio às atividades de pesquisa das instituições. Buscando a interoperabilidade entre os sistemas CRIS, a maioria está baseada no formato europeu Common European Research Information Format (CERIF), que fornece um modelo de dados apresentado na Figura 1, do qual podem ser utilizadas todas ou algumas das suas entidades e relações de acordo com as necessidades da instituição.

Figura 1 - Modelo de dados CERIF



Fonte: Common European Research Information Format (2013).

Desta forma, entre os serviços que podem ser oferecidos a partir da informação contida em um ambiente CRIS e das funções/aplicações que ele tem, são apresentados os seguintes serviços (ZIMMERMAN, 2002; euroCRIS, 2014; CASTRO; SHEARER; SUMMANN, 2014; SAYÃO; SALES, 2015):

a) gestão integrada dos dados, informações e o ciclo de vida da pesquisa;
b) geração de relatórios e informações confiáveis sobre índices e resultados das pesquisas para tomada de decisões sobre o processo de pesquisa;
c) intercâmbio de informação, conhecimentos, equipamentos, infraestrutura e comunicação entre cientistas;
d) acesso a informações necessárias para a elaboração de políticas baseadas em dados atuais e alocação de recursos ou fontes de financiamento;
e) avaliação de projetos e resultados para transferência de tecnologia e a detecção e análise de tendências nas ciências;
f) identificação de novos mercados para produtos de pesquisa e de concorrência e potenciais parceiros.

Nesta lista pode-se notar que existem vários serviços que um CRIS oferece a partir da integração dos dados de pesquisa. Entre eles, destaca-se a produção

de relatórios que permitirão identificação de pesquisadores, equipamentos e infraestrutura especializada para evitar a duplicação de esforços e resultados. A referida produção permite, ainda, a troca de conhecimento e a avaliação dos projetos, que levará ao reconhecimento de resultados de pesquisa, a futuras fontes de financiamento e a novos mercados e políticas de investigação, entre outros, de acordo com as informações contidas em um determinado tipo de CRIS e das necessidades informacionais e organizacionais da instituição.

5 Discussão: análise da natureza dos serviços na integração de RD e CRIS

Os RD e os CRIS são soluções cujo objetivo comum é a gestão de informação relacionada com a pesquisa científica, tanto o armazenamento de dados e/ou resultados do processo de investigação para acesso livre (publicações e *datasets*, dentre outros serviços realizados pelos repositórios), quanto com o registro dos metadados relacionados com as atividades do ciclo de vida da pesquisa (preocupação dos CRIS).

Ademais, existem claras diferenças entre estes tipos de sistemas, como se observa no Quadro 1, em que Sales e Sayão (2015) resumem e apresentam uma adaptação do exposto por Castro, Shearer e Summann (2014) a este respeito.

Quadro 1 - Principais diferenças entre sistemas CRIS e RI

(continua)

QUESTÃO	SISTEMAS	
	CRIS	RI
Função principal	Gestão da informação sobre o ciclo total das atividades de pesquisa científica.	Formação da memória acadêmica digital das instituições de pesquisa.
Escopo	Lida com a informação de todo o ciclo de atividade de pesquisa, incluindo informações sobre projetos, financiamentos, pessoas, organizações, produtos de pesquisa, instalações e equipamentos.	Ênfase no arquivamento de produtos de pesquisa, principalmente publicações acadêmicas. Entretanto, o escopo tem se estendido para incluir outros produtos, tais como dados de pesquisa.

Quadro 1 - Principais diferenças entre sistemas CRIS e RI

(continuação)

Objetivo	Coletar informação com o objetivo de reportar as atividades institucionais de pesquisa, considerando também informação sobre ambientes externos como governo e agências de fomento à pesquisa.	Arquivar, tratar, preservar e disseminar produtos de pesquisa, com forte ênfase em publicações acadêmicas, tendo como orientação os pressupostos do movimento de Livre Acesso.
Abrangência	Na qualidade de ferramenta voltada para coletar informação sobre todas as áreas da instituição envolvidas em atividades de pesquisa, incluindo informações econômicas, os sistemas CRIS são basicamente projetados para uso interno.	Os repositórios por sua vez, são orientados para o mundo exterior, ou seja, disseminar, dar visibilidade, permitir o acesso livre e aumentar o impacto dos produtos de pesquisa da instituição.
Gestão	Os sistemas CRIS são tipicamente comandados pelos mesmos setores que fazem a gestão dos projetos de pesquisa na instituição e que cuidam da avaliação, planejamento, finanças etc.	Os repositórios são geralmente gerenciados pelas unidades de informação das instituições, representadas, na maioria dos casos, pelas bibliotecas de pesquisa.
Conteúdo	Como ferramenta acurada para reportar atividades de pesquisa, os sistemas CRIS lidam com metadados que contextualizam os fluxos de pesquisa da instituição.	Os repositórios também fazem uso de metadados, porém sua atenção está centrada na disponibilidade de texto completo e na criação, arquivamento, curadoria e reuso dos objetos digitais.

Quadro 1 - Principais diferenças entre sistemas CRIS e RI

(conclusão)

Metadados	Os sistemas CRIS operam com padrões de metadados complexos e não suficientemente harmonizados, mas que são necessários para a descrição das inúmeras atividades de pesquisa. O CERIF é o padrão mais usado. Embora a sua implementação varie consideravelmente entre países, ele está se tornando-se rapidamente o padrão default que poderá contribuir para a interoperabilidade de alto nível.	Os repositórios operam com modelos bem mais simples de metadados, que são, por sua vez, transversalmente bastante consistentes entre instituições e países, isto permite um alto grau de interoperabilidade. O padrão mais usado é o Dublin Core qualificado.
Plataformas de software	As instituições desenvolvem suas próprias soluções de CRIS no sentido de atender aos seus requisitos específicos e igualmente as soluções comerciais são utilizadas principalmente devido à complexidade intrínseca dos modelos e à possibilidade de configurar as comunidades de usuários.	As plataformas de software para Repositórios têm uma arquitetura mais simples e são, tradicionalmente, pacotes do tipo *open source* que envolvem, no seu desenvolvimento, uma ampla comunidade internacional. Os pacotes mais usados mundialmente são o DSpace e o EPrints.
Compartilhamento de informações	As comunidades CRIS só recentemente iniciaram o compartilhamento sistemático de informações. Ainda incipiente, devido à questão de metadados, modelos adotados e padronização.	As comunidades responsáveis pelos repositórios, orientadas pelo padrão de intercâmbio bibliotecas de pesquisa, praticam uma troca intensiva de informações.
Interoperabilidade	Ainda incipiente, devido à questão de metadados, modelos adotados e padronização.	Via protocolo OAI-PMH e o uso de metadados DC.

Fonte: Sales e Sayão (2015, p. 176).

As diferenças entre esses sistemas podem estar relacionadas à evolução independente de cada um e, especialmente, aos objetivos que definiram sua criação. Entretanto autores como Castro, Shearer e Summann (2014) afirmam que a interoperabilidade entre sistemas CRIS e RD é generalizada e permitirá o intercâmbio de informação eficiente entre essas plataformas, reforçando as características de cada um delas.

Em relação à integração RD/CRIS, encontram-se várias configurações ou arquiteturas para lidar com características e/ou funcionalidades de ambos os sistemas em uma instituição, que permitirão a gestão eficiente da informação de pesquisa como um todo. A determinação de uma configuração específica dependerá de vários fatores, tais como:

> [...] o tamanho da instituição, os relatórios de que precisa, recursos de TI disponíveis para atualização e manutenção do sistema ou o equilíbrio de poder entre as diferentes unidades institucionais em determinado momento. (CASTRO, 2014, s/p).

A partir desses fatores, as políticas e as necessidades da instituição e o RD e/ou CRIS existente(s) são apresentados no Quadro 2, uma síntese das configurações que podem ser implementadas para fornecer serviços da integração RD/CRIS por Castro, Shearer e Summann (2014).

Quadro 2 - Configurações para gestão da informação de RD/CRIS

(continua)

Configuração	Descrição	Exemplos
RD (com características de CRIS)	O modelo de dados de um repositório institucional existente se estende adicionando-lhe uma estrutura de metadados que descrevem outros elementos relacionados com os objetos armazenados lá e que fazem parte do ciclo de vida de pesquisa tais como projetos, financiamentos, etc. Esses metadados são uma melhoria específica para o repositório que adiciona características e / ou funcionalidades do sistema CRIS.	DSpace-CRIS: é uma solução que permite que um RD possa gerenciar informação que tradicionalmente estaria em CRIS, como pessoas e projetos, entre outros.

Quadro 2 - Configurações para gestão da informação de RD/CRIS

(conclusão)

RD + CRIS	O RD e CRIS são plataformas que operam de forma independente em uma instituição apoiando os processos de investigação separadamente, pela eficiência e com a disponibilidade tecnológica nesta configuração os dois sistemas são interoperáveis através de uma troca automática de dados que normalmente é feito com o mapeamento CERIF (CRIS) /DC ou MODS (IR) e com o apoio de protocolos de comunicação.	Complemento CRIS para EPrints: é uma atualização que adiciona ao modelo de dados do repositório, caraterísticas de CRIS.
CRIS (com características de RD)	No caso de plataformas CRIS comerciais ou desenvolvidas pela instituição, elas executaram funções de RD que permitem o armazenamento e publicação de objetos sem a existência de um repositório próprio e independente. Ou seja, a informação relacionada com a investigação e gestão dos resultados da investigação (próprio de um RD) são feitas pelo CRIS.	Diversas características podem ser agregadas em repositórios para auxiliar no povoamento de registros no repositório e ainda auxiliar na alimentação de bases de dados de autores como ORCID, assim como a inclusão de número de citações de bases de dados, a inclusão de métricas alternativas (Altmetrics), são claros elementos de um CRIS que podem ser adicionados a um RD. CRIS baseado CERIF + RD baseado em Dublin Core.

Fonte: Adaptado de Castro, Shearer e Summann (2014).

Cada instituição deverá avaliar, a partir da existência de um ou ambos os sistemas, quais dessas configurações se adaptam à sua realidade, para fins diferentes e complementares, pelas particularidades ou fatores antes mencionados e o âmbito da informação de pesquisa que se pretende gerenciar.

Com respeito à interoperabilidade RD/CRIS, ela "[...] redimensiona o valor desses sistemas na medida em que seus ativos informacionais se tornam matéria prima para serviços inéditos, de maior alcance em termos de usuários" (SALES; SAYÃO, 2015, p. 154), ou seja, há uma oportunidade latente de prestação de serviços decorrentes da interoperabilidade RD/CRIS, produto da informação que gera cada um, o que permitirá uma maior cobertura das necessidades informacionais dos usuários desses sistemas.

De acordo com Castro (2014), pode-se dizer também que a integração das funcionalidades de RD e CRIS leva a uma complementação desses sistemas como um todo e a sua pontencialização, porquanto essa integração visa melhorar as características individuais de cada um e melhor servir as necessidades de uma instituição, sendo isto conseguido por meio da interoperabilidade entre as plataformas.

Analisando o que foi exposto sobre sistemas de informação e serviços informacionais, Yurchyshyna *et al*. (2011, p. 265) relatam que:

> [...] é importante distinguir dois tipos de iniciativas, definidas pela sua origem: (i) iniciativa como parte dos sistemas de informação; (ii) iniciativa como um serviço informativo. Uma iniciativa que faz parte dos sistemas de informação, que já existem e funcionam, tem por objetivo melhorar e manter os serviços existentes, integrando-os e, como tal, melhorar as relações entre os atores envolvidos (...) Uma iniciativa que é considerada como um serviço informativo visa a criação de novos serviços. O apoio a tais iniciativas é um dos principais elementos de uma abordagem inovadora para criar valor através da informação.

Além disso, os mesmos autores afirmam que existe uma natureza complementar entre o impacto nos serviços com respeito aos dois tipos de iniciativas anteriores:

> Em correspondência com dois tipos principais de iniciativas, podemos estabelecer dois tipos principais de seu impacto nos serviços. (...) i) a evolução dos serviços - quando definem os

> princípios da integração e do posicionamento dos serviços nos sistemas de informação existentes (...); e (ii) a inovação de serviços - quando ajudam a identificar conhecimentos que poderiam tornar-se passíveis de ação e, como tal, levariam à criação de novos serviços. (...) os processos de inovação e evolução dos serviços são complementares. (YURCHYSHYNA *et al*., 2011, p. 266).

Uma análise da integração RD/CRIS com foco nessa natureza complementar permite visá-la como uma iniciativa com origem nos ambientes informacionais RD e CRIS, ou seja, seu impacto é uma evolução de serviços a partir da existência em cada uma dessas plataformas que já têm estabelecidos os seus próprios serviços.

Uma vez estabelecido e compreendido que a integração ou interoperabilidade RD/CRIS reforça cada um desses e pode ser considerada como uma proposta de evolução dos serviços existentes em cada uma das plataformas e não como inovação de serviços, o Quadro 3 apresenta oportunidades de melhoria que podem ser consideradas para essa evolução a partir da integração discutida.

Quadro 3 - Oportunidades para evolução de serviços na integração RD/CRIS

ID	OPORTUNIDADE
1	Otimizar a visibilidade dos documentos e do repositório.
2	Gerenciar a disponibilidade e restrição de acesso aberto (embargos, etc).
3	Considerar identificadores persistentes, estatísticas de uso e valores bibliométricos.
4	Prever a preservação de longo prazo.
5	Gerir publicações.
6	Disponibilizar interfaces (APIs, HTTP, REST, SRU, OAI-PMH).
7	Melhorar a usabilidade do usuário final (busca, incorporação).
8	Permitir o processamento de metadados de fornecedores externos.
9	Garantir a qualidade dos metadados e atividades de preservação dos dados.
10	Possibilitar a integração com sistemas externos, como plataformas de gestão de dados de pesquisa, ambientes virtuais de pesquisa, plataformas de publicação, como sistemas de periódicos e CRIS.
11	Disponibilizar o suporte ao estilo de citação.
12	Implementar o suporte de *linked open data*.

Fonte: Adaptado de Castro, Shearer e Summann (2014)

Os diversos serviços que podem ser adicionados com intuito de apresentar melhorias na interação do cliente com a interface do Repositório ou do CRIS, por exemplo, a última oportunidade do Quadro 3, "Implementar o suporte de *Linked Open Data*", está atrelada à estruturação dos dados no contexto do *Linked Data* (dados ligados). Segundo Arakaki (2016, p. 119), o *Linked Data* está relacionado "[...] às melhores práticas para estruturação dos dados a partir das ferramentas [da Web Semântica] e a garantia de troca de informações." Uma aplicação do *Linked Data* na integração RD/CRIS, conforme o mesmo autor:

> [...] amplia as possibilidades de navegação do recurso informacional no ambiente da Web, permitindo a identificação de diversos outros recursos relacionados ao que se está visualizando. Logo, o usuário poderá otimizar a forma de navegar (...) podendo escolher um determinado recurso informacional, suas derivações e seus recursos relacionados. (ARAKAKI, 2016, p. 15).

Nessa perspectiva, os usuários poderão navegar pelo ambiente fornecido na integração do RD/CRIS e localizar os recursos informacionais por meio de seus relacionamentos. Destaca-se, entretanto, que isso só poderá ser feito com a estruturação dos dados combinada com as tecnologias da Web Semântica.

Finalmente, dada a complementaridade entre a evolução e a inovação dos serviços, as oportunidades listadas podem também ser analisadas no foco da criação de novos serviços pela integração de RD/CRIS.

6 Considerações finais

Neste trabalho, a preocupação foi a análise das iniciativas de serviços em repositórios digitais e CRIS. A partir deste estudo, observou-se que os serviços nesses ambientes tendem a uma integração dos sistemas para ampliar as possibilidades de serviços oferecidos.

Nas bibliotecas e centros de pesquisa, o gerenciamento de informação relacionado com as atividades de pesquisa, incluindo os produtos finais, pode

abranger a interoperabilidade de repositórios digitais de registros de artigos publicados em periódicos, teses, dissertações, patentes, dados científicos, entre outros. Enquanto que, os *Current Research Information System* armazenam metadados dos fluxos de trabalho dessas atividades (entidades e relacionamentos), além de informações dos pesquisadores, suas redes de colaboração, etc.

Foram encontradas diversas configurações para integração de RD/CRIS, conforme a política adotada pela instituição para o armazenamento e disseminação da informação. Além disso, vem sendo fortalecido o desenvolvimento dessa integração que conduz para reforçar as características de cada uma das plataformas na organização, preservação, recuperação, acesso e disseminação de seus conteúdos para o público geral ou um público restrito.

A Ciência de Serviços constitui-se num auxílio à explicação e ao aprimoramento das interações entre várias entidades para alcançar melhores resultados para o benefício mútuo com a integração RD/CRIS. O seu impacto pode ser considerado como uma evolução de serviços, permitindo potencializar oportunidades relacionadas a: otimização da visibilidade do RD e do CRIS, gestão de publicações, usabilidade do usuário final, processamento de metadados de fornecedores externos (elementos de CRIS), suporte de *Linked Open Data*, entre outros.

Finalmente, no foco da natureza complementar existente entre a evolução e a inovação de serviços, a criação de novos serviços poderá ser concretizada a partir das análises das necessidades informacionais da instituição e de seus usuários com a da iniciativa da integração das plataformas RD/CRIS nos serviços informativos.

Referências

ARAKAKI, F. A. **Linked data**: ligação de dados bibliográficos. 2016. 144 f. Dissertação (Mestrado em Ciência da Informação) - Universidade Estadual Paulista, 2016. Disponível em: https://repositorio.unesp.br/bitstream/handle/11449/147979/arakaki_fa_me_mar.pdf?sequence=2&isAllowed=y. Acesso em: 29 jun. 2018.

BARROS, M. H. T. C. de. **Disseminação da informação**: entre a teoria e a prática. Marília: [s. n.], 2003.

CALABRIA, P. C. *et al*. A ciência da inovação em serviços: estudo exploratório sobre os interesses e prioridades para uma agenda de pesquisa no Brasil. **Revista de Administração e Inovação**, v. 10, n. 4, p. 110-135, 2013.

CASTRO, P. 7 things you should know about…institutional repositories, CRIS systems, and their Interoperability. *In*: REPOSITORY OBSERVATORY, 3., 2014, s. l. **Tópicos** [...] Hamburg, 2014. Disponível em: https://www.coar-repositories.org/community/events/archive/repository-observatory-third-edition/7-things-you-should-know-about-institutional-repositories-cris-systems-and-their-interoperability/. Acesso em: 13 jun. 2018.

CASTRO, P. de; SHEARER, K.; SUMMANN, Friedrich. The gradual merging of repository and CRIS solutions to meet institutional research information management requirements. **Procedia Computer Science**, v. 33, p. 39-46, 2014. DOI: https://doi.org/10.1016/j.procs.2014.06.007. Disponível em: https://www.sciencedirect.com/science/article/pii/S1877050914007972?via%3Dihub. Acesso em: 10 jun. 2018.

CHERSBROUGH, H. W. A failing grade for the innovation academy. **Financial times** (North American edition), 2004. Disponível em: http://www.ft.com/cms/s/2/9b743b2a-0e0b-11d9-97d3-00000e2511c8.html#axzz2nBjCpSlj. Acesso: 13 jun. 2018.

CHESBROUGH, H. W.; SPOHRER, J. I. A research manifesto for services science. **Communications of the ACM**, v. 49, n., p. 35-40, jul. 2006. DOI: 10.1145/1139922.1139945. Disponível em: http://dl.acm.org/citation.cfm?id=1139945. Acesso em: 13 jun. 2018.

COMMON EUROPEAN RESEARCH INFORMATION FORMAT. **CERIF-1.6**. 2013. Disponível em: http://www.eurocris.org/cerif/feature-tour/cerif-16. Acesso em: 13 jun. 2018.

DHOLAKIA, N.; MUNDORF, N.; DHOLAKIA, R. R. Novos serviços de informação e comunicação: um quadro de referência estratégico. **Ciência da Informação**, v. 26, n. 3, p. 1-8, set./dez. 1997. DOI: http://dx.doi.org/10.1590/S0100-19651997000300002. Disponível em: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-19651997000300002&lng=pt&tlng=pt. Acesso em: 7 jun. 2018.

euroCRIS. **CRIS concept and CRIS benefits**. 2010. Disponível em: http://bit.ly/2uj45HL. Acesso em: 19 mar. 2018.

FERRARI, R. D. D.; PIRES, G. L. F. A. Auto-arquivamento e acesso aberto: deveres e direitos digitais na sociedade em rede. **Revista Digital de Biblioteconomia e Ciência da Informação**, v. 12, n. 1, p. 22-38, jan./abr. 2014. DOI: https://doi.org/10.20396/rdbci.v12i1.1616. Disponível em: https://periodicos.sbu.unicamp.br/ojs/index.php/rdbci/article/view/1616/pdf_48. Acesso em: 13 jun. 2018.

FERREIRA, S. M. S. P. Repositórios versus revistas científicas: convergências e convivências. *In*: FERREIRA, S. M. S. P.; TARGINO, M. das G. (org.). **Mais sobre revistas científicas**: em foco a gestão. São Paulo: Editora Senac São Paulo, 2008. p. 111-137.

GIANESI, I. G. N.; CORRÊA, H. L. **Administração estratégica de serviços**: operações para a satisfação do cliente. São Paulo: Atlas, 1994.

KATZ, M. L.; SHJAPIRO, C. Technology adoption in the presence of network externalities. **Journal of Political Economy**, v. 94, n. 4, p. 822–841, ago. 1986. DOI: https://doi.org/10.1086/261409. Disponível em: http://www.journals.uchicago.edu/doi/10.1086/261409. Acesso em: 13 jun. 2018.

KURAMOTO, Hélio. Informação científica: proposta de um novo modelo para o Brasil. **Ciência da Informação**, Brasília, v. 35, n. 2, p. 91-102, maio/ago 2006.

LEITE, F. C. L. **Como gerenciar e ampliar a visibilidade da informação científica brasileira**: repositórios institucionais de acesso aberto. Brasília: IBICT, 2009.

LOVELOCK, C. H.; WRIGHT, L. **Serviços**: marketing e gestão. São Paulo: Saraiva, 2001.

MAGLIO, P. P. *et al*. The service system is the basic abstraction of service science. **Information Systems and e-Business Management**, v. 7, p. 395–406, set. 2009. DOI: https://doi.org/10.1007/s10257-008-0105-1. Disponível em: https://link.springer.com/article/10.1007/s10257-008-0105-1. Acesso em: 1 jun. 2018.

PINTO, C. S.; SIMÕES, C.; AMARAL, L. CERIF: is the standard helping to improve CRIS? **Procedia Computer Science**, v. 33, p. 80-85, 2014. DOI: http://dx.doi.org/10.1016/j.procs.2014.06.013. Disponível em: https://www.sciencedirect.com/science/article/pii/S1877050914008035?via%3Dihub. Acesso em: 8 jun. 2018

SALES, L. F.; SAYÃO, L. F. Ciberinfraestrutura de informação para a pesquisa: uma proposta de arquitetura para integração de repositórios e sistemas CRIS. **Informação & Sociedade**, v. 25, n. 3, p. 163-184, set./dez. 2015. Disponível em: http://www.periodicos.ufpb.br/ojs2/index.php/ies/article/view/23998. Acesso em: 13 jun. 2018.

SANTAREM SEGUNDO, J. E. **Representação iterativa**: um modelo para repositórios digitais. 2010. Tese (Doutorado em Ciência da Informação) – Universidade Estadual Paulista "Júlio de Mesquita Filho", Marília, 2010. Disponível em: http://hdl.handle.net/11449/103346. Acesso em: 13 jun. 2018.

SANTOS, L. C.; FACHIN, G. R. B.; VARVAKIS, G. Gerenciando processos de serviços em bibliotecas. **Ciência da Informação**, v. 32, n. 2, p. 85-94, maio/ago. 2003. Disponível em: http://revista.ibict.br/ciinf/article/view/1009/1064. Acesso em: 13 jun. 2018.

SPOHRER, J.; KWAN, S.K. Service Science, Management, Engineering, and Design (SSMED): An Emerging Discipline - Outline and References. **International Journal of Information Systems in the Service Sector**, v. 1, n. 3, 2009.

SPOHRER, J.; MAGLIO, P. P. Service science: toward a smart planet. In: KARWOWSKI, Waldemar; SALVENDY, Gavriel. **Introduction to service engineering**. New York: Wiley & Sons, 2010. p. 3-30.

SPOHRER, J.; MAGLIO, P. P. The emergence of service science: toward systematic service innovations to accelerate co-creation of value. **Production and Operations Management**, v. 17, n. 1, p. 238-246, 2008. DOI: https://doi.org/10.3401/poms.1080.0027. Disponível em: http://onlinelibrary.wiley.com/doi/10.3401/poms.1080.0027/abstract. Acesso em: 13 jun. 2018.

VIDOTTI, S. A. B. G. *et al*. Coleta automática para povoamento de repositórios digitais: conversão de registros utilizando XSLT. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 17., 2016, Salvador. **Anais** [...]. Salvador: UFBA, 2016. Disponível em: http://www.ufpb.br/evento/index.php/enancib2016/enancib2016/paper/viewFile/4092/2552. Acesso em: 13 jun. 2018.

YURCHYSHYNA, A. *et al*. Innovation and Evolution of Services: Role of Initiatives. *In*: TAMBOURIS E.; MACINTOSH A.; BRUIJN, H. de. Electronic Participation: ePart. **Lecture Notes in Computer Science**. Berlim: Springer, 2011. v. 6847, p. 262-273. Disponível em: https://link.springer.com/chapter/10.1007%2F978-3-642-23333-3_23. Acesso em: 13 jun. 2018.

ZIMMERMAN, E. CRIS-Cross: Current Research Information Systems at a Crossroads. In: INTERNATIONAL CONFERENCE ON CURRENT RESEARCH INFORMATION SYSTEMS, 6., 2002, Kassel. **Trabalhos apresentados** [...] Kassel: University of Kassel, 2002. p. 12-21.

Ecologia informacional complexa das bibliotecas universitárias

Edgar Bisset Alvarez
Silvana Aparecida Borsetti Gregorio Vidotti
Juan Antonio Pastor-Sánchez

1 Introdução

O papel das Tecnologias de Informação e Comunicação (TIC) nas bibliotecas universitárias, junto ao desenvolvimento e quebra de paradigmas do ensino superior, converteu esta instituição, ao longo dos anos, em um componente imprescindível aos processos docente-educativos nas universidades, sendo "[...] um sistema de informação que é parte de um sistema mais amplo, que poderia ser chamado de sistema de informação acadêmico, no qual, a geração de conhecimentos é o objeto da vida universitária" (FUJITA, 2005, p. 1-2).

As Bibliotecas Universitárias são consideradas o maior espaço e ambiente informacional de uma universidade, geralmente com presença nos meios digital e analógico., constituindo-se como a maior referência no que diz respeito à busca, acesso, disseminação, organização, armazenamento e gestão da informação. Suas salas e espaços são os ambientes onde os usuários costumam, de forma geral, acessar a informação da qual precisam, contando para isso com serviços, recursos e produtos informacionais vinculados a processos que permitem cumprir seus objetivos.

Entretanto, como bem expressa Fujita (2005, p. 7), "[...] o momento atual vivido pelos Sistemas de Informação exige o estabelecimento de um Programa de Gestão de Informação e de Pessoas, que considere a cultura da Instituição acadêmica, inserindo-a na Sociedade da Informação, resultado da Globalização vivida pela humanidade".

Nesse sentido, é preciso pensar na relação que se estabelece entre pessoas, processos, tecnologias e espaços de informação, para poder gerenciar melhor os diferentes e cada vez mais complexos ambientes de informação que surgem como resultado do acelerado avanço tecnológico. Esta relação, metaforicamente,

Davenport e Pruzak (1998) chamaram de "Ecologia Informacional", descrevendo-a como "[...] a administração holística da informação ou administração informacional centrada no ser humano." (DAVENPORT; PRUZAK, 1998, p.18).

No entanto, com o acelerado avanço das tecnologias, foi necessário o desenvolvimento de ações que visassem organizar, disponibilizar e tornar acessível toda a informação disponível nos diferentes ambientes informacionais, principalmente, sob a ótica de levar o usuário ao encontro da informação precisa, no momento exato, com a maior qualidade e o menor esforço por parte dele. Esta afirmação condizente com uma das principais funções das Bibliotecas Universitárias, fica fortalecida nas palavras de Wersig e Nevelling (1975, p. 11) quando da necessidade de "[...] transmitir o conhecimento para aqueles que dele necessitam".

Atualmente, considerando o alto grau de desenvolvimento tecnológico alcançado pela sociedade, o acesso, o uso, o gerenciamento e o compartilhamento da informação, as Bibliotecas Universitárias concentram-se na solução dos problemas informacionais que o indivíduo socialmente reconhecido possa apresentar em um contexto e momento determinado.

Focado nessa linha de trabalho, Bisset (2017) apresenta um modelo de Arquitetura da Informação para Bibliotecas Universitárias, visando trabalhar na construção de ambientes informacionais cada vez mais abertos, em que usuários possam ter acesso aos recursos, produtos e serviços de informação da universidade sem restrições e caraterizados principalmente pelo seu dinamismo, integração com os processos universitários e capacidade para fomentar o intercâmbio de conhecimento entre os usuários e propiciar a aprendizagem.

Para alcançar seu objetivo, Bisset (2017) utiliza um modelo de Arquitetura da Informação Pervasiva proposto por Oliveira (2014) que visava, essencialmente, melhorar os ambientes informacionais para garantir que o usuário desenvolva uma experiência mais atrativa a cada interação, tendo como premissa central "[...] que se algo não for encontrado, então não será usado. Ou, inversamente, quanto mais encontrável uma coisa é, muito mais provável é que seja usada." (WILKIE; AZZOPARDI, 2013, p. 808, tradução nossa).

Potencializa-se assim, nas palavras de Vechiatto, Oliveira e Vidotti (2016), a preocupação com inúmeros ambientes informacionais, com informações relevantes para os usuários, mas deficientes na projeção de interfaces voltadas às experiências e competências dos sujeitos Informacionais. Portanto as Bibliotecas Universitárias, consideradas um serviço de suporte e apoio para a universidade, diante do avanço vertiginoso das tecnologias e o surgimento de sistemas e ambientes cada vez mais complexos, estão obrigadas a construir espaços e entornos informacionais que estejam atrelados às principais funções e áreas de atuação.

2 Bibliotecas universitárias: ambientes, serviços e perfis

As Bibliotecas Universitárias surgem, pela primeira vez, na idade média, em pleno século XII, vinculadas ao surgimento das universidades como centros de ensino, tendo no início como principal função a proteção do patrimônio documental da humanidade, para oferecê-lo a seus usuários, principalmente estudantes e professores, com o intuito de oferecer uma formação prática no uso de livros que os capacitavam para exercer uma profissão e dela garantir seu sustento.

É definida pela Americam Library Association como "[...] uma combinação orgânica de pessoas, coleções e instalações, cujo propósito é ajudar a seus usuários no processo de transformar a informação em conhecimento [...]" (AMERICAN LIBRARY ASSOCIATION, 1989, p. 2, tradução nossa).

As bibliotecas universitárias sempre tiveram um forte vínculo com o desenvolvimento das universidades e desempenharam um importante papel para o avanço da sociedade. Fazendo um breve resumo da vastíssima história destas instituições, Vianna (2013, p. 12, grifo nosso) traz a sua visão da evolução histórica das Bibliotecas Universitárias e a divide em três grandes períodos:

> • **Do século XII ao início do século XX - Biblioteca Tradicional**: compreende da origem das universidades (1080) até o começo do processo de automação (~1960/70), com foco no acervo.
> • **Até o final do século XX - Biblioteca Automatizada**: com foco na redução do trabalho manual, através do

> desenvolvimento de software para a localização e o empréstimo do acervo.
> • **A partir do século XXI - Biblioteca ubíqua** e de uso autônomo: destaca-se pelo desenvolvimento de software para adquirir, localizar, emprestar e acessar a informação de forma remota, focada na informação, no acesso on-line e na autonomia do usuário.

Desde o começo do século XXI, uma das principais preocupações no contexto biblioteconômico foram as mudanças na interação "biblioteca-usuário". A ampla conectividade que o desenvolvimento tecnológico começava a oferecer e a grande quantidade de conteúdos criados, manipulados e disseminados só favoreciam e ainda favorecem o aumento da preocupação das instituições de informação (em especial as bibliotecas) com a criação e desenvolvimento de serviços, recursos e produtos de informação para satisfazer o usuário, cada vez mais conectado e independente.

Neste contexto, as bibliotecas universitárias têm sido obrigadas a assumir um papel mais proativo na sua relação com o público, papel este que Orera-Orera (2007, p. 5) destacou quando se referia a "[...] um novo modelo de biblioteca para um novo modelo de sociedade [...]", advertindo que as mudanças a serem adotadas não poderiam ser superficiais.

Na tentativa de alcançar uma sistematização, no que diz respeito ao futuro das bibliotecas universitárias, encontraram-se vários posicionamentos, que diante das referidas mudanças elucidam o papel, função e objetivos das bibliotecas universitárias, resumidos no seguinte posicionamento:

> Espaços físicos e virtuais, flexíveis, onde convergem e se integram infraestruturas tecnológicas, recursos humanos, espaços, equipamentos e serviços de informação (proporcionados em qualquer momento e acessíveis desde qualquer lugar) orientados à pesquisa e à aprendizagem. (DOMÍNGUEZ, 2005).

Porém a biblioteca universitária é desafiada a lidar com uma tríade de ambientes: "o analógico tradicional", "o digital ou virtual" e "o híbrido". Nesse sentido, o uso das tecnologias da informação e comunicação marca um papel icônico nas mudanças dos ambientes de informação das bibliotecas universitárias:

> 1. A criação e desenvolvimento de espaços virtuais de aprendizagem, para a formação dos usuários,
> 2. O aumento na criação e uso de novos serviços e recursos de informação (Bases de dados, bibliotecas digitais, repositórios de informação, etc.),
> 3. Aumento na acessibilidade dos serviços e produtos de informação da biblioteca de forma online. (BICET, 2012, p. 33).

Estas mudanças nos contextos das bibliotecas universitárias também guardam uma estreita relação com as mudanças das universidades e seu contínuo processo de automatização, bem como com o surgimento de um novo conceito de universidade, a "universidade virtual", como afirmaram Medina e Vicente (2008, p. 4): "As implicações, no campo da informação, são de enorme transcendência para as bibliotecas, que encontram nesta necessidade, um novo impulso e alento a seu papel como expertas na gestão da informação".

2.1 Novos ambientes de informação digital

O surgimento da Internet e o desenvolvimento e posterior evolução da World Wide Web atuaram como motor principal das mais significativas transformações dos ambientes digitais, em especial, dos ambientes informacionais digitais no contexto biblioteconômico.

O uso desta plataforma universal como ponte para dar acesso a seus serviços, recursos de informação e disponibilizar novos ambientes informacionais - bibliotecas digitais, repositórios institucionais, catálogos online, etc. – demonstra a importância dada pelas bibliotecas ao fato de que a maioria dos processos de informação de seus usuários começava a desenvolver-se no contexto digital.

Em um primeiro momento, o desenvolvimento da web caracterizou-se pela existência de ambientes de informação estáticos e de pouca interatividade. Esta plataforma universal foi desenhada, em seu início, principalmente, com a finalidade de propiciar apenas o consumo de informação pelos humanos, em um contexto, onde os sistemas criados não se comunicavam nem se entendiam.

Geralmente, nesse período, as bibliotecas universitárias limitavam-se a usar seus ambientes de informação digital para disponibilizar páginas web com serviços e produtos de informação, além de notícias de interesse dos seus usuários, isso no que diz respeito ao funcionamento das bibliotecas. Logo este espaço constituía-se como uma via para oferecer um serviço unidirecional e de pouca interação entre Biblioteca-Usuário.

Já no começo do século XXI, os avanços nas TIC seguiram aumentando e as bibliotecas tentando acompanhar este processo evolutivo. Nesta relação Bibliotecas-TIC tem-se como segundo momento de mudança a introdução de aplicações, tecnologias e serviços, tendo como objetivo o aumento da interatividade da web por meio de uma maior participação e utilização da inteligência coletiva, por parte dos usuários, no uso, criação e consumo da informação na web. Este fenômeno deu nome a um novo conceito de web, a "Web 2.0", termo mencionado pela primeira vez pela "O'Reilly Media" no ano de 2004 em uma série de conferências com MediaLive International (O´RELLY, 2005, p. 1).

Diante desta situação, as bibliotecas universitárias reagiram, modificando muitos dos seus processos e serviços, assim como a forma como interagiam com seus usuários, convertendo seus sites e ambientes de informação digitais em ambientes mais colaborativos e participativos, a partir do uso das tecnologias e serviços da web 2.0 (wikis, blogs, canais RSS, redes sociais, chats entre outras). Fenômeno este usado para abrir o debate sobre um novo modelo de biblioteca, a "Biblioteca 2.0".

Conforme Medina e Vicente (2008, p. 3, tradução nossa):

> Este conceito de web implica uma profunda mudança de atitude no que diz respeito à informação, na forma em que é concebida e utilizada, o que supõe um paradigma absolutamente diferente

> e novas expectativas em relação à tecnologia. A informação que se gera nestas redes sociais compreende também os ambientes de pesquisa e aprendizagem e está inter-relacionada com a rede social global da web.

Outro elemento a ser considerado na evolução da web e sua relação com a introdução das TIC nas bibliotecas universitárias, refere-se à proposta de Berners-Lee, Hendler e Lassila (2001), de construção de uma web onde a informação interligada semanticamente fosse seu eixo principal de funcionamento. Tal web foi denominada "Web Semântica", e foi pensada como uma extensão da web atual, na qual a informação se posiciona com significado bem definido e estruturado, que permite sua compreensão pelos computadores, facilitando assim uma relação de cooperação entre ser humano e computador, em que a informação legível por máquina representa a ação das pessoas e o sistema informático faz a análise dos padrões de interação.

Esta nova proposta apresenta para as bibliotecas universitárias um novo usuário, a "Máquina". E neste contexto, estas instituições que até então se preocupavam apenas com a informação gerada, usada e compartilhada pelo usuário humano, passam a objetivar e dotar de significado os conteúdos da web para as máquinas, o que permitiria a estas usar e criar informações a partir de sua inter-relação com humanos ou com outras máquinas. Este fenômeno exige das bibliotecas universitárias uma maior coerência em suas arquiteturas e na concepção de seus ambientes de informação digital, visando criar sistemas e serviços de informação compreensíveis e utilizáveis não só por humanos, mas também pelas máquinas.

Como se demonstrou até agora, a internet teve um forte impacto no desenvolvimento dos serviços, ambientes e produtos de informação das bibliotecas universitárias, mas ela não pode ser considerada o único fator importante neste processo. Embora o uso da internet tenha, de fato, aumentado e amplificado a acessibilidade dos serviços das bibliotecas universitárias, os processos bibliotecários, por sua vez, davam seus primeiros passos no âmbito digital, a partir do desenvolvimento de sistemas que permitiam o gerenciamento dos

processos e fluxos de informação, que antigamente estavam estreitamente ligados ao contexto analógico.

Uma das primeiras incursões das TIC no contexto bibliotecário foi a automatização dos catálogos, com a criação do formato para descrição bibliográfica Machine Readable Cataloging (MARC). Inicia-se, então, um longo caminho na automatização dos diferentes processos bibliotecários, sendo a busca e recuperação da informação um dos processos centrais que acompanharam as primeiras experiências das bibliotecas, na introdução das TIC em seu contexto.

O surgimento dos registros legíveis por máquinas foi seguido pelo aparecimento dos primeiros recursos eletrônicos e documentos digitais oriundos dos processos de digitalização iniciados nas últimas décadas do século passado.

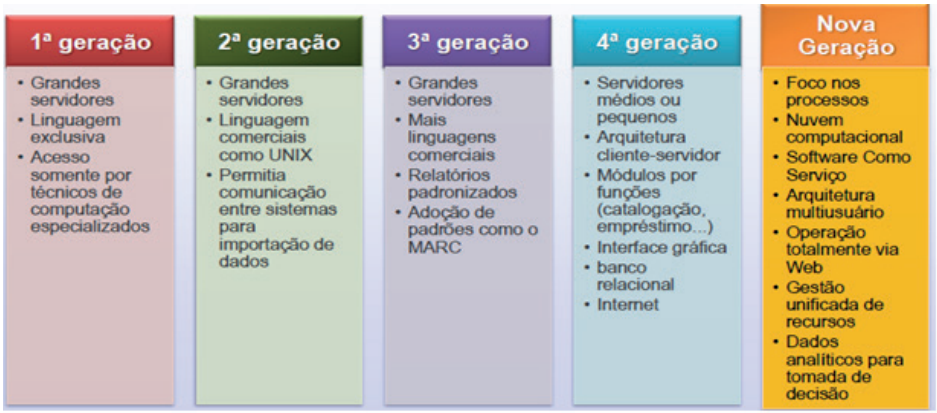
Posteriormente, com a evolução e o surgimento de novas tecnologias, a importância que fora outorgada à informação e aos processos de comunicação promoveu o desenvolvimento de sistemas de informação automatizados, destinados a apoiar os bibliotecários nos processos de gestão, administração, organização, recuperação e disponibilização da informação.

Nascem e evoluem, assim, diferentes sistemas informacionais, que de variadas formas contribuíram para o processo de automação das bibliotecas, são eles: os Sistemas de Gerenciamento de Bibliotecas: dirigidos a gestão informática e automatização das diferentes atividades existentes dentro de uma biblioteca (gestão de coleções e dos usuários, catalogação, aquisições, consultas e estatísticas, etc.); os Repositórios Digitais: dedicados a manutenção, proteção e preservação de informação em formato digital, sendo o tipo mais conhecido o "repositório acadêmico ou institucional"; as Bibliotecas Digitais: dedicadas a organização, preservação e tratamento técnico de coleções de objetos digitais, bem como a sua disponibilização para o acesso; os Sistemas de Gerenciamento de Revistas Eletrônicas: criados para a administração online de revistas, geralmente de caráter científico; os Sistemas de busca e Descoberta: com a função de recuperar, integrar e apresentar em uma só interface os conteúdos e os documentos disponíveis em diferentes fontes de informação, sejam elas remotas ou locais.

Na Figura 1, Vianna (2014, p. 20) representa a evolução dos sistemas usados pelas bibliotecas no processo de automação de seus processos e serviços

informacionais, separando-os em cinco gerações que acompanham os avanços que a revolução tecnológica das últimas décadas tem imposto ao desenvolvimento da sociedade: as gerações de sistemas para bibliotecas.

Figura 1 - As gerações de sistemas para bibliotecas



Fonte: Vianna (2014).

Por outro lado, o contínuo avanço das tecnologias e a variedade de sistemas que convergem entre si e que são introduzidos no contexto bibliotecário e, em especial, nas bibliotecas universitárias, trazem consigo outro aspecto tecnológico no qual os profissionais da informação e as bibliotecas devem prestar muita atenção, pois a construção dos ambientes de informação digitais torna-se um elemento de suma importância para o cumprimento dos objetivos das bibliotecas universitárias.

Prática esta que se converteu em um elemento primordial para as bibliotecas universitárias e para os profissionais que nelas trabalham, através da criação e elaboração de projetos de arquitetura da informação que façam uso de aplicações mais dinâmicas e de fácil acesso, facilitando a interoperabilidade entre os sistemas de uma mesma biblioteca, assim como com o de outras bibliotecas, o que evitaria o surgimento de ambientes de informação que se assemelhem a verdadeiras caixas pretas ou silos de informação.

2.2 Novos serviços

Durante séculos, as bibliotecas universitárias tiveram como função principal e final a proteção e conservação do patrimônio cultural e intelectual da humanidade. A partir do fim do século XIX, estas instituições têm sido marcadas pela abertura de suas coleções e desde os 80, no século XX, pela consolidação dos conceitos de usuários e serviços de informação, assim como pelo aumento da necessidade de dirigir as funções e serviços das bibliotecas, tendo como embasamento científico um processo de matematização do conhecimento científico, fatos estes que levaram a uma mudança na filosofia de trabalho das bibliotecas universitárias.

Hoje, estas instituições são consideradas peças-chave no desenvolvimento científico, social, político, econômico e cultural de qualquer país. O aumento desenfreado da produção documental, principalmente no ambiente digital, o avanço das TIC e sua incursão no contexto das bibliotecas, além de destacar o protagonismo dos usuários no processo de criação, gerenciamento, uso e compartilhamento da informação, constituíram-se em fatores responsáveis pela conversão das bibliotecas universitárias em agentes de mudança, com a missão de, através de seus serviços, converterem-se em centros de reencontro, debate e reflexão entre aqueles envolvidos, ou não, no convívio universitário, assim como de ser a principal ferramenta na formação e autoaprendizagem de seus usuários.

Este processo de reinvenção da biblioteca universitária passa pela criação de serviços de informação, contextualizados com a realidade daqueles que destes fazem uso. Azevedo (2001, p. 1) identifica os serviços de informação como "[...] Atividade destinada à identificação, aquisição, processamento e transmissão de informação ou dados e ao seu fornecimento num produto de informação [...]".

De acordo com o dicionário Houaiss (2016), a palavra serviço, faz referência a "[...] ação ou efeito de servir, de dar de si algo em forma de trabalho [...]" ou um "[...] produto da atividade humana destinada à satisfação de necessidades, mas que não apresenta o aspecto de um bem material [...]", por sua vez, Lancaster (1979, p. 13, tradução nossa) descreve que "A informação é uma coisa bastante intangível, não é possível vê-la, ouvi-la ou senti-la. Estamos informados no que

diz respeito a uma matéria, se o nosso estado de conhecimento sobre aquela matéria se transforma de alguma forma".

Seguindo nesta linha de pensamento, as bibliotecas universitárias podem ser consideradas instituições fornecedoras de serviços, em específico, de serviço de informação, uma vez que se utilizam de produtos e ferramentas (bases de dados, aplicações tecnológicas, publicações etc.) para satisfazer as necessidades de informação de seus usuários e assim mediar o processo de apropriação da informação.

Segundo Lancaster (1996), os serviços em uma biblioteca podem ser divididos em dois grandes grupos: o primeiro é constituído por aqueles serviços voltados para as atividades internas da instituição (chamados de serviços técnicos), o segundo é constituído por aqueles serviços voltados para a comunidade que a biblioteca serve (chamados de serviços ao público) e é neste último grupo que a presente pesquisa concentra seu foco.

Quando se fala de serviços de informação nas bibliotecas universitárias, é preciso fazer uma distinção entre os chamados serviços tradicionais, aos quais estas bibliotecas dedicaram-se a vida inteira, e entre aqueles que são produto da evolução e dos avanços científicos e tecnológicos, bem como das mudanças originadas no contexto do ensino superior e que surgem como resposta às necessidades de uma comunidade usuária mais participativa e ao mesmo tempo mais independente.

Ao falar-se de serviços tradicionais de informação, ou seja, daqueles serviços que sempre formaram e formam ainda hoje parte da vida das bibliotecas, em específico das universitárias, pode-se citar a consulta de documentos em sala, empréstimo, serviço de referência e informação bibliográfica, reprografia, empréstimo entre bibliotecas, disseminação seletiva da informação, formação de usuários, etc.

Por outro lado, o desenvolvimento acelerado das TIC, a influência de uma economia totalmente globalizada, o aumento do investimento nas universidades na década dos 1990, a dedicação de parte desse orçamento ao desenvolvimento e atualização das bibliotecas universitárias, bem como a mudança de paradigma no ensino superior, fizeram com que as bibliotecas universitárias deixassem de

funcionar isoladamente e começassem a trabalhar coletivamente, formando as redes de bibliotecas universitárias.

Nesse sentido, as bibliotecas universitárias começam um novo ciclo, modificando seu estilo de trabalho, seus serviços e concentrando seu foco no usuário final, demonstrando assim uma nova concepção de biblioteca universitária, caracterizada por:

> [...] a) oferecer atenção personalizada e responder às necessidades dos usuários, sejam estes docentes, pesquisadores ou estudantes; b) coordenar e otimizar todo o conjunto de recursos e serviços que existem dispersados pela universidade e que são potencialmente necessários e úteis para o desenvolvimento das atividades da docência, aprendizado e pesquisa; c) usar as tecnologias da informação e comunicação como instrumentos para a acessibilidade e organização da informação; e d) assumir funções alfabetizadoras dos membros da comunidade universitária. (AREA-MOREIRA, 2004, p. 73, tradução nossa).

Neste contexto, os serviços de informação estão voltados a satisfazer um usuário com um comportamento informacional que se tornou diferente, deixando de ser passivo, um mero receptor, para ser ativo, criador, gerenciador e distribuidor da informação, obrigando, por consequência, as bibliotecas a oferecerem e criarem serviços de informação em ambientes de informação físicos e digitais, que representem as mais variadas necessidades e interesses, o que, por sua vez, demanda arquiteturas da informação mais colaborativas e democráticas.

Assim, Sunyer (2006), Celestino (2007) e Pinto, Sales e Osorio (2008) oferecem um resumo daqueles serviços de informação que, a seu ver, devem estar presentes no contexto atual das Bibliotecas Universitárias:

- Acesso à informação científica de qualidade e de acesso aberto. (Sunyer, Celestino, Pinto, Sales y Osorio).

- Recursos de informação específicos para melhorar ou adquirir competências.
- Formação em habilidades informacionais.
- Orientação na criação de conteúdos de docência e pesquisa.
- Consultoria em aspectos vinculados com a proteção legal da informação.
- Suporte no desenvolvimento de congressos e seminários.
- Serviços de acolhimento, informação geral e informação especializada.
- Serviço de criação de materiais multimídia. (Celestino (2007), Pinto, Sales e Osorio (2008)).
- Serviço de criação de materiais impressos.
- Serviço de laboratório de línguas. (Celestino (2007), Pinto, Sales e Osorio (2008)).
- Serviço de gravação de projetos audiovisuais.
- Serviço de suporte informático.
- Serviço de apoio à inovação docente e ao aprendizado.
- Serviço de salas de estudo, de trabalho em grupos, seminários. (BICET, 2012, p. 35).

Um elemento a destacar-se na relação apresentada pelos autores acima é o nível de participação e colaboração do usuário e a concentração dos serviços de informação na sua figura, aumentando o fluxo e o intercâmbio de informação usuário-biblioteca (feedback), o que, sem dúvida, facilita os processos de avaliação, criação e implantação de serviços de informação, a partir de uma filosofia de arquitetura da participação exposta por O´Relly (2005) e que, na sua essência,

propõe a participação ativa do usuário nos processos de construção e organização dos conteúdos.

Aos poucos, as bibliotecas universitárias vêm se transformado e adaptando-se às novas exigências da sociedade contemporânea, pois, durante muito tempo, os serviços de informação prestados por estas instituições foram baseados principalmente na única tarefa de dar acesso aos documentos armazenados em suas coleções.

Então, o que mudou? Pode-se dizer que, com o decorrer do tempo, mudaram a forma de acesso e a diversidade de novos serviços criados a partir da introdução das tecnologias da informação. Serviços estes convergentes e integrados que inter-relacionam todos os aspectos que incidem na vida acadêmica, formativa e de pesquisa de professores e estudantes.

2.3 Perfil de uso das bibliotecas universitárias

Todas as mudanças pelas quais têm passado as bibliotecas universitárias são resultado de um processo de reestruturação e realinhamento na sua filosofia de trabalho, e que tem como principal objetivo atualizar seus processos, serviços e técnicas para adaptar-se a um novo momento histórico, em que a automação dos serviços e processos informacionais tem aumentado o fluxo de informações, facilitando o intercâmbio nas relações interpessoais e a valorização da informação e do conhecimento científico-técnico.

Esta seção trata de ilustrar como estão sendo usadas, de forma geral, as bibliotecas universitárias em função de todas as mudanças que nelas estão acontecendo.

Quando consultada a bibliografia que aborda os estudos de usuários em bibliotecas universitárias, encontram-se, de forma geral, 04 categorias de usuários com as quais esta instituição tem trabalhado ao longo do tempo: 1- Professores e Pesquisadores; 2- Alunos de Graduação e Pós-Graduação; 3- Pessoal administrativo; e 4- Visitantes ou Externos. E se o objetivo é analisar o perfil de uso das bibliotecas universitárias na atualidade, resulta inevitável discorrer sobre quais são as características dos usuários destas instituições atualmente.

Desde o começo do século XXI, Prensky (2001) já demonstrava certa preocupação com as características dos usuários que faziam uso dos ambientes de informação digital das bibliotecas universitárias, apresentando uma classificação que os distinguia em dois tipos de usuários: "Digital Immigrants" e "Digital Natives". O primeiro grupo faz referência àqueles usuários nascidos na era analógica e que vêm se adaptando aos desafios que o contexto digital lhes apresenta, e o segundo grupo se refere a todos aqueles usuários já nascidos no contexto digital, que por este motivo assimilam muito mais facilmente qualquer mudança de natureza tecnológica.

A existência destas duas tipologias de usuários, categorizadas em função de suas habilidades tecnológicas, resulta em um caminho para as bibliotecas universitárias, pois a substituição do analógico pelo digital traz consigo, de forma implícita, uma mudança na mentalidade daqueles imigrantes digitais, em comparação com os nativos digitais, pois a forma com que os primeiros estavam acostumados a consumir informação mudou radicalmente, levando as bibliotecas universitárias a desenvolver ações para formar habilidades nestes usuários no uso das TIC, ao mesmo tempo em que implanta processos e serviços cada dia mais automatizados:

> A maioria dos estudantes que iniciam o ensino superior hoje em dia é muito mais jovem que o microcomputador, e se sentem mais confortáveis trabalhando em um teclado que escrevendo em um caderno espiralado e são mais felizes lendo na tela do computador que em um papel na mão. A conexão contínua (estando em contato com amigos e familiais em qualquer momento e desde qualquer lugar) é de suma importância para eles. (FRAND, 2000, p. 15, tradução nossa).

No entanto, os imigrantes digitais, assim como os nativos digitais, são considerados pelas bibliotecas universitárias usuários virtuais, pois, no entendimento de Manso-Rodríguez (2007, p. 4), estas duas categorias de usuários apresentam uma série de elementos em comum: "[...] a forma remota com que acessam a informação;

não têm acesso fácil aos documentos físicos, é difícil o processo de traslado até a instituição, necessitam rapidez e imediatismo e valorizam a personalização [...]", semelhanças estas nas quais as bibliotecas vêm apostando para o desenho e implementação de serviços em ambientes virtuais, cada vez mais personalizados, com ambientes de informação e interfaces mais amigáveis e acessíveis ou, como bem aponta Margaix (2007, p. 101),"[...] serviços atrativos, que sejam úteis para os usuários e lhes convidem a participar e a desenvolver seu conhecimento".

> [...] seria um erro acreditar que só a busca de informação dos estudantes foi modificada pela opção digital massiva, o acesso contínuo (24 horas/7 dias) ao material científico, a eliminação da intermediação pelos poderosos e influentes motores de busca também foram. O mesmo tem ocorrido com professores, conferencistas e profissionais. Todos demonstram um comportamento vital e rápido, que os leva a procurar mais horizontal que verticalmente [...]" (BRITISH LIBRARY; JOINT INFORMATION SYSTEMS COMMITTEE, 2008, p. 8, tradução nossa).

As bibliotecas universitárias possuidoras de uma tradição de grandes coleções impressas e históricas enfrentam hoje seu maior desafio, um mercado digital em constante crescimento e usuários que a cada dia ficam mais afastados da biblioteca, enquanto ambiente físico, e demandam, cada vez mais, conteúdos dinâmicos e personalizados, livre acesso 24 horas por dia e 07 dias por semana, além de respostas em um *click*, o que obriga estas instituições a entrar na competição pela atenção desses usuários.

Além das habilidades no uso das TIC, o foco das bibliotecas universitárias nos interesses e comportamentos dos usuários leva ao crescimento de ambientes com ofertas informativas que realmente representam as necessidades e interesses da sua comunidade usuária.

Neste processo de direcionar mais atenção ao usuário, ouvi-lo, encorajá-lo a participar na construção dos serviços, processos e produtos que ele mesmo irá

usar, aumenta, consideravelmente, o processo de feedback usuário-biblioteca, colaborando no aumento da qualidade dos serviços, na construção das coleções, na representação e organização da informação, assim como na co-participação para solução das necessidades de outros usuários.

Por outro lado, as bibliotecas universitárias vêm trabalhando também na formação de profissionais da informação cada vez mais motivadores, humanos, responsáveis, criativos, inovadores, pesquisadores, formadores e mediadores no processo de apropriação da informação pelos usuários.

Neste mundo onde prevalece a informação em formato digital, caracterizado pela forma de acesso massivo e fácil aos conteúdos disponíveis na web, as informações disponibilizadas pelas bibliotecas universitárias apresentam-se com acesso cada vez mais difícil se comparado com as facilidades oferecidas pelas ferramentas de busca como o Google e outros motores de busca, e isso é preocupante, pois já é fato que a geração atual de consumidores da informação não demonstra grande interesse em comprovar a veracidade da informação recuperada.

A *British Library* e o *Joint Information Systems Committee* (JISC) (2007), analisando o informe elaborado pela *Online Computer Library Center* (OCLC), apresentam o comportamento dos usuários referente ao uso dos motores de busca em comparação com o uso das bibliotecas, demonstrando que:

> • Aproximadamente 90% dos estudantes universitários fazem uso de motores de busca para começar um processo de busca de informação e somente 2% começam a partir da página web da biblioteca;
>
> • Aproximadamente 93% sentem-se satisfeitos ou muito satisfeitos com sua experiência, de forma geral, no uso dos motores de busca e 84% com a busca assistida por um bibliotecário;
>
> • Os motores de busca adequam-se melhor aos estilos de vida dos estudantes universitários do que as bibliotecas físicas ou online;

> • Os estudantes universitários ainda usam a biblioteca, mas com menor frequência (leem menos), desde que começaram a usar as ferramentas de pesquisa da Internet;
>
> • Para esta geração, os livros constituem a principal imagem associada à biblioteca, apesar do massivo investimento que estas instituições fazem em recursos digitais. (BRITISH LIBRARY; JOINT INFORMATION SYSTEMS COMMITTEE, 2008, p. 7, tradução nossa).

Estes dados demonstram como os usuários têm preferência pelo uso dos motores de busca para o processo de busca, recuperação e acesso à informação, devido à rapidez e globalidade que estas ferramentas oferecem, em comparação com as buscas sofisticadas e complexas oferecidas pelos sistemas de informação das bibliotecas universitárias. Bastos (2013, p. 249) conclui em seu estudo sobre "A interação do usuário com catálogos bibliográficos on-line: investigação a partir da Teoria Fundamentada" que:

> [...] a preferência da maioria dos usuários ao interagir com o catálogo durante seu processo de busca, demonstra que eles optam por uma busca rápida, adotando a opção de busca padrão do catálogo e de forma anônima, ou seja, sem se identificarem no sistema, e os dados evidenciam que apesar dos recursos disponíveis nos catálogos bibliográficos, em função dos avanços tecnológicos, os usuários não utilizam os serviços oferecidos.

Outro estudo da OCLC 2010 aponta que os estudantes das universidades:

> • (43%) sentem que fontes de informação das bibliotecas universitárias são mais confiáveis que as dos motores de busca;
>
> • (57%) são mais propensos a utilizar o sítio web da biblioteca;
>
> • (92%) são mais propensos a usar páginas de redes sociais;

> • (83%) começam sua busca de informação online, utilizando motores de busca;
>
> • (88%) usam Wikipédia. (BRITISH LIBRARY; JOINT INFORMATION SYSTEMS COMMITTEE, 2008, p .8, tradução nossa).

Estes resultados demonstram que os usuários das bibliotecas universitárias estão utilizando ferramentas que demandam poucas habilidades e que são simples de manusear no processo de busca. Isso explica por que os usuários de hoje em dia, segundo a British Library e Joint Information Systems Committee (2008, p. 7, tradução nossa), realizam as suas pesquisas sem levar em conta a estrutura da biblioteca ou a forma na qual a mesma segmenta os diferentes recursos nas distintas áreas da sua página web.

São esses mesmos usuários que hoje demandam das bibliotecas serviços de informação mais dinâmicos, imediatistas e adaptados à sua realidade, uma realidade marcada pelo uso intensivo de dispositivos móveis, pelo armazenamento em nuvem e pela exploração massiva das redes sociais para a criação, produção, consumo, uso e compartilhamento da informação, fato este que desconstrói a diferenciação entre consumidor e produtor de informação.

No entanto, as bibliotecas universitárias seguem destacando-se por sempre estar à frente na aplicação dos avanços tecnológicos em seus processos e serviços de informação, o que unido às novas demandas dos usuários favorece uma nova forma de orientar estes serviços e produtos, em função das necessidades da sua comunidade usuária, fazendo deles serviços mais interativos e colaborativos e, sobretudo, buscando antecipar-se às necessidades dos usuários, oferecendo informações em forma de recomendações.

3 Ecologia informacional das bibliotecas universitárias

Quando se revisa a literatura em torno da evolução das Bibliotecas Universitárias, pode-se observar a existência de cinco elementos que a maioria

dos autores deixam claro no que diz respeito ao comportamento das bibliotecas universitárias e as caraterísticas que as distinguem como ambientes informacionais:

a) as coleções;

b) os serviços;

c) os profissionais da informação;

d) a sua vinculação às instituições universitárias;

e) os usuários, seu contexto e relações sociais que influenciam no seu comportamento informacional.

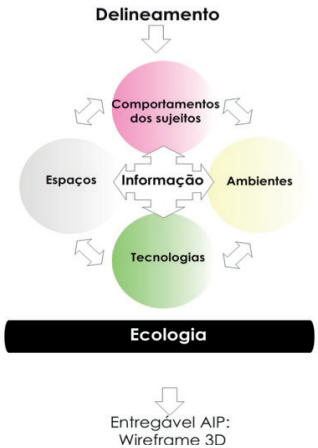
É por isso que trabalhar na construção de ambientes informacionais de Bibliotecas Universitárias na atualidade pode ser considerado um desafio para qualquer profissional de informação, principalmente se busca-se garantir o hibridismo dos seus ambientes informacionais e o uso de instrumentos e tecnologias, em função de melhorar a experiência e o comportamento informacional dos usuários, considerando a informação um elemento transversal a todos os processos e com a premissa de que esta esteja disponível para o usuário, independentemente do ambiente, contexto ou dispositivo que esteja utilizando.

Dito isto, pode-se afirmar que estamos lidando com um modelo de Bibliotecas Universitárias, do qual a comunidade acadêmica espera:

> Informação **on-demand**;
>
> Informação **just-in-time**;
>
> **Informação** pelo seu **conteúdo**, não pelo seu suporte;
>
> Acesso **ubíquo**;
>
> **Autonomia** no acesso à informação;
>
> Sistemas de **metabusca** e **descoberta** unificados;
>
> Acesso por quaisquer **dispositivos fixos ou móveis**. (VIANA, 2014, p. 8, grifos do autor).

Estas características, quase que intrínsecas ao novo modelo de biblioteca universitária, permite-nos considerá-lo uma Ecologia Informacional Complexa, descrita por Oliveira (2014, p. 135) como aquela que "[...] integra holisticamente espaços, ambientes, tecnologias e os sujeitos com seus comportamentos através da informação.", como se demonstra na Figura 2, e que ao mesmo tempo constitui objeto/fenômeno de investigação da Arquitetura da Informação Pervasiva.

Figura 2 - Estrutura básica das Ecologias Informacionais Complexas



Fonte: Oliveira (2014).

Neste sentido, Oliveira (2014, p. 135) assegura que "[...] as ações de natureza prática estão relacionadas ao projeto, acompanhamento e avaliação de ecologias informacionais complexas em contextos da realidade pós-moderna [...]". Considerado isto, Vechiato, Oliveira e Vidotti (2016, p. 3762) afirmam que "No contexto da Arquitetura da Informação Pervasiva, a Encontrabilidade da Informação se posiciona como um fim, como um objetivo geral a ser atingido no projeto, avaliação ou implementação de ecologias informacionais complexas [...]". Por outro lado, Oliveira (2014) afirma que em uma ecologia informacional complexa, a encontrabilidade da informação propicia a interação dos sujeitos com os conjuntos de dados e de informações nela contidos.

Segundo Davenport e Pruzak (1998, p. 14), "[...] A ecologia da informação exige novas estruturas administrativas, incentivos e atitudes em direção à hierarquia, à complexidade e à divisão de recursos da organização [...]", descrevendo-a como "[...] a ciência de compreender e administrar todos os ambientes [...]".

Por outro lado, Morville (2004, p. 2) destaca que para que um determinado sistema ou ambiente permita uma experiência agradável deve existir uma estreita e forte relação entre os elementos, contexto, conteúdo, usuário, pois, isto outorgaria certa autonomia aos usuários em sua interação nos ambientes de informação. O autor, quando aborda o tema, fá-lo desde um contexto digital, mas cada uma dessas facetas poderia ser aplicada também a ambientes analógicos ou híbridos.

As bibliotecas universitárias, neste contexto tecnológico em que a sociedade vive, buscam respostas e instrumentos que lhes permitam viabilizar uma gestão eficaz e eficiente do grande volume de informações que armazenam, principalmente em função de melhorar a experiência informacional dos usuários nos seus ambientes.

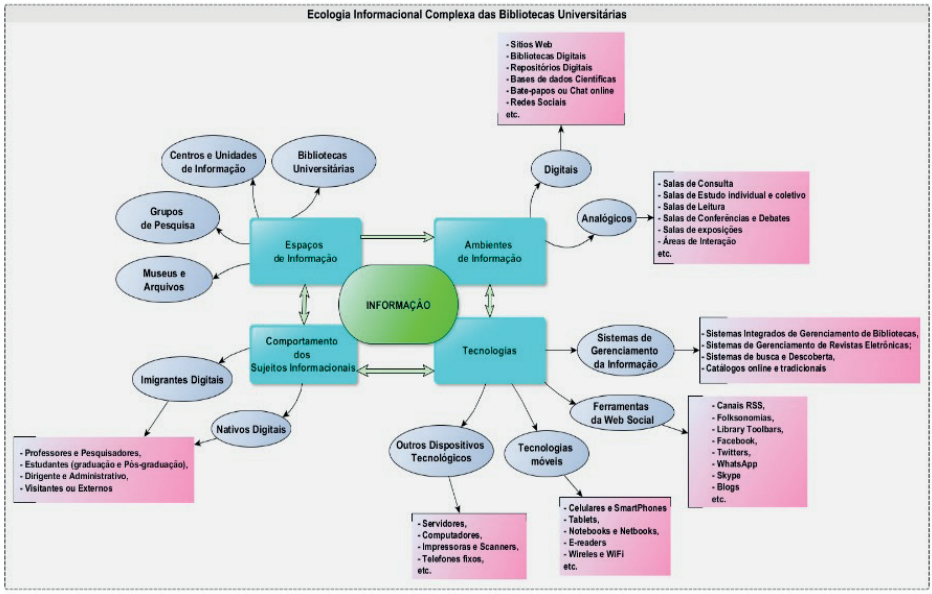
Logo é fato que existe certa preocupação pela construção de ambientes de informação, mais flexíveis e abertos, com estruturas mais horizontais e modulares, que permitam intercambiar informação entre elas, facilitando a criação de serviços mais personalizados, ou, como sugerem os autores Resmini e Rosati (2011, p. 189, tradução nossa), "Não se concentrar apenas nas relações hierárquicas entre itens [...] e [...] fortalecer as relações horizontais, assim como as interligadas pela similaridade, acoplamento ou comportamento social"[3].

Sendo assim, e tendo como elemento pautador a construção de ecologias informacionais, bem como considerando as perspectivas de Davenport e Pruzak (1998) e de Oliveira (2014), pode-se determinar os componentes que formam parte das ecologias informacionais das atuais Bibliotecas Universitárias (tecnologias, espaços e ambientes informacionais, as necessidades e comportamento dos usuários, os dispositivos tecnológicos).

3 "[...] do not focus only on the hierarchical relationships between ítems [...] e [...] strengthen horizontal relationships such as those implied by similarity, coupling, or social behavior [...]"

Na Figura 3, apresenta-se a ecologia informacional complexa de uma Biblioteca Universitária, trazendo uma visualização das principais categorias (Espaços Informacionais, Ambientes de Informação, Tecnologias, Comportamentos dos Sujeitos Informacionais), "[...] devolvendo o homem ao centro do mundo da informação." e facilitando aos "[...] ecologistas da informação podem mobilizar não apenas designs arquiteturais e TI, mas também estratégia, política e comportamento ligados à informação, além de suporte a equipes e processos de trabalho para produzir ambientes informacionais melhores [...]" (DAVENPORT; PRUZAK, 1998, p. 21).

Figura 3 - Ecologia Informacional Complexa das Bibliotecas Universitárias



Fonte: Autoria própria (2018).

As Bibliotecas Universitárias, de forma geral, possuem espaços e ambientes informacionais tradicionais, digitais e/ou híbridos para a prestação dos serviços que ela oferece. O usuário, fazendo uso das tecnologias, pode buscar, recuperar, acessar, usar e compartilhar informações. Isto permitiu fornecer uma representação

estruturada e claramente explícita das Ecologias Informacionais das Bibliotecas Universitárias, vinculando os múltiplos sistemas de informação, para garantir o apoio aos processos docentes, investigativos e culturais das universidades.

4 Considerações finais

As transformações que a sociedade vem sofrendo, produto dos avanços nas áreas das tecnologias, desde meados do século XX, momento histórico-social que esteve marcado por grandes descobertas científicas e transformações tecnológicas que ajudaram o homem a alcançar um maior conhecimento do mundo ao seu redor e a modificar sua forma de interagir com ele, significou a passagem de uma sociedade industrial para uma sociedade altamente dependente da informação.

Nesse sentido, a Biblioteca Universitária precisa criar ambientes onde exista uma maior integração e interoperabilidade entre seus fluxos e sistemas de informação digitais, analógicos e/ou híbridos e onde o usuário possa ter as mesmas experiências informacionais, indiferente do contexto digital, analógico ou híbrido.

Nesta direção, estão sendo focadas todas as mudanças que no contexto das bibliotecas universitárias têm se produzido nas últimas décadas, tendo como objetivo final atender a um universo de usuários com interesses, necessidades, costumes e, geralmente, com hábitos tecnológicos muito diferentes e instáveis.

Na construção de uma ecologia informacional, visa-se mapear os ambientes e espaços informacionais e identificar os recursos informativos, as necessidades informacionais e Tecnologias que procuram propiciar experiências positivas, combinando a proatividade do ambiente e os sistemas com elementos como usabilidade e acessibilidade, levando os usuários a descobrir as informações que podem satisfazer suas necessidades, facilitando os processos de encontro, uso, busca, recuperação e gerenciamento da informação.

Diante deste cenário e tendo como substrato teórico as abordagens para a construção das Ecologias Informacionais, apresentou-se uma representação que permite a visualização, mapeamento, categorização e organização dos elementos essenciais que compõem toda ecologia, para assim ter uma melhor compreensão

de quais são as ações a serem tomadas durante a elaboração e construção da Ecologia Informacional de uma Biblioteca Universitária.

Referências

AMERICAN LIBRARY ASSOCIATION. Standards for university libraries: evaluation of performance. *In*: **American Library Association** (Ed.). 1989.

AREA-MOREIRA, M. **De la Biblioteca Universitaria al Centro de Recursos para el Aprendizaje e Investigación**: elaboración de uma guía sobre la organización y gestión de un crai en el contexto de las universidades españolas. Informe Final sobre el papel que habrían de jugar las bibliotecas universitarias en el nuevo modelo de enseñanza aprendizaje que se deriva de la implantación de los denominados "créditos europeos" (ECTS), 2004. Disponível em: http://www.ub.edu/esbrina/docs/area-crai.pdf. Acesso em: 28 fev. 2018.

AZEVEDO, A. **Serviço de informação**. 2001. 35 transparências, color. Disponível em: https://paginas.fe.up.pt/~fsilva/mgi/files/Servicos_de_informacao.ppt. Acesso em: 1 jun. 2018.

BASTOS, F. M. **A interação do usuário com catálogos bibliográficos on-line**: investigação a partir da teoria fundamentada. 2013. Tese (Doutorado em Ciência da Informação) – Universidade Estadual Paulista "Júlio de Mesquita Filho", Marília, 2013.

BERNERS-LEE, T.; HENDLER, J.; LASSILA, O. The semantic web: a new form of web content that is meaningful to computers will unleash a revolution of new possibilities. **Scientific American Special**, número especial, 2001. Disponível em: https://www-sop.inria.fr/acacia/cours/essi2006/Scientific%20American_%20Feature%20Article_%20The%20Semantic%20Web_%20May%202001.pdf. Acesso em: 25 abr. 2018.

BICET, E. **Propuesta de pautas para el diseño de un centro de recursos para el aprendizaje y la investigación como modelo de trabajo para la Red de Bibliotecas de la Universidad de la Habana**. 2012. Dissertação (Mestrado em Biblioteconomia e Ciência da Informação) – Universidade da Havana, Havana, 2012.

BISSET, E. **Sistemas de recomendação para bibliotecas universitárias**: um aporte teórico da arquitetura da informação. 2017. Tese (Doutorado em Ciência da Informação) – Universidade Estadual Paulista "Júlio de Mesquita Filho", Marília, 2017. Disponível em: https://repositorio.unesp.br/handle/11449/151685. Acesso em: 15 jun. 2018.

BRITISH LIBRARY; JOINT INFORMATION SYSTEMS COMMITTEE. **Information behaviour of the researcher of the future**. 2008. Disponível em: http://www.webarchive.org.uk/wayback/archive/20140614113419/http://www.jisc.ac.uk/media/documents/programmes/reppres/gg_final_keynote_11012008.pdf. Acesso em: 30 mar. 2018.

CELESTINO, S. Los CRAI en Europa: cambio cultura y estratégias en la gestión de recursos humanos. Em Jornada: **Profesionales para los CRAI ¿estamos preparados para los cambios?** Madrid, 2007.

DAVENPORT, T. H; PRUSAK, L. **Ecologia da informação**. São Paulo: Futura, 1998.

DOMINGUEZ, M. T. S. La función social de las bibliotecas universitarias. **Boletín de la Asociación Andaluza de Bibliotecarios**, n. 80, p. 43-70, 2005.

FRAND, J. The information-age mindset: changes in students and implications for higher education. **Educause Review**, p. 15-24, mar./abr. 2000. Disponível em: https://er.educause.edu/-/media/files/article-downloads/erm0051.pdf. Acessado em: 03 jan. 2018.

FUJITA, M. S. L. A biblioteca digital no contexto da gestão de bibliotecas universitárias: análise de aspectos conceituais e evolutivos para a organização da informação. *In*: ENCONTRO NACIONAL DE CIÊNCIA DA INFORMAÇÃO, 6., 2005, Salvador. **Anais** [...] Salvador: UFBA, 2005. Disponível em: http://www.cinform-anteriores.ufba.br/vi_anais/docs/MariangelaFujita.pdf. Acesso em: 02 dez. 2018.

HOUAISS, A. Dicionário Houaiss da língua portuguesa. 2016. Disponível em: https://houaiss.uol.com.br/. Acesso em 02 dez. 2018.

LANCASTER, F. W. **Avaliação de serviços de bibliotecas**. Brasília, DF: Briquet de Lemos, 1996.

LANCASTER, F. W. **Information retrieval systems**: characteristics, testing and evaluation. 2.ed. New York: J. Wiley, 1979. 381p.

MANSO-RODRÍGUEZ, R. A. **Bibliotecas, bibliotecarios y usuarios**: una aproximación a los desafíos y oportunidades en un mundo 2.0. La Habana, 2007.

MARGAIX, D. A. Conceptos de web 2.0 y biblioteca 2.0: origen, definiciones y retos para las bibliotecas actuales. **El profesional de la información**, v. 16, n. 2, p. 95-106, mar./abr. 2007. Disponível em: http://eprints.rclis.org/9521/1/kx5j65q110j51203.pdf. Acesso em: 30 mar. 2018.

MEDINA, A. L.; VICENTE, L. Z. Las bibliotecas universitarias y la gestión de la información en el entorno digital: unas consideraciones para repositorios digitales. **BiD: textos universitaris de biblioteconomia i documentació**, n. 20, jun. 2008. Disponível em: http://bid.ub.edu/20lopez2.htm. Acesso em: 10 maio 2018.

MORVILLE, P. **User experience design**. 2004. Disponível em: http://semanticstudios.com/publications/semantics/000029.php. Acesso em: 30 mar. 2018.

O'REILLY, T. **What is Web 2.0**. 2005. Disponível em: http://www.oreilly.com/go/web2. Acesso em: 29 maio 2018.

OLIVEIRA, H. P. C. de. **Arquitetura da informação pervasiva**: contribuições conceituais. 2014. Tese (Doutorado em Ciência da Informação) – Universidade Estadual Paulista "Júlio de Mesquita Filho", Marília, 2014.

ORERA-ORERA, L. La biblioteca universitaria ante el nuevo modelo social y educativo. **El profesional de la información**, v. 16, n. 4, p. 329-337, 2007.

PINTO, M.; SALES, D.; OSORIO, P. **Biblioteca universitaria, CRAI y alfabetización informacional**. Gijón: Ediciones Trea, 2008.

PRENSKY, M. Digital natives, digital immigrants. **On the Horizon**, v. 9, n. 5, out. 2001. Disponível em: http://www.marcprensky.com/writing/Prensky%20-%20Digital%20Natives,%20Digital%20Immigrants%20-%20Part1.pdf. Acesso em: 29 mar. 2018.

RESMINI, A.; ROSATI, L. **Pervasive information architecture**: designing cross-channel user experiences. Burlington: Elsevier, 2011.

SUNYER, S. Centros de recursos para el aprendizaje y la investigación - CRAI -, y servicios bibliotecarios estratégicos para una Europa basada en el conocimiento. **Intangible Capital**, v. 2, n. 14, p. 327-337, 2006.

VECHIATO, F. L.; OLIVEIRA, H. P. C.; VIDOTTI, S. A. B. G. Arquitetura da informação pervasiva e encontrabilidade da informação: instrumento para a avaliação de ambientes informacionais híbridos. *In*: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 17., 2016, Salvador. **Anais** [...]. Salvador: UFBA, 2016.

VIANA, M. Sistemas de informação para bibliotecas universitárias: análise multicritérios. *In*: **Seminário Nacional de Bibliotecas Universitárias**, 18., – Minicurso, 2014.

WERSIG, G., NEVELlNG, U. The phenomena of interest to information science. **Information Scientist**, v. 9, p. 127-140, 1975.

WILKIE C., AZZOPARDI L. An initial investigation on the relationship between usage and findability. *In*: SERDYUKOV, P. et al. (ed.). Advances in information retrieval. **Lecture Notes in Computer Science**. Berlim: Springer, 2013. v. 7814.

Perspectivas de investigação em encontrabilidade da informação: em foco o atributo Descoberta de Informação

Fernando Luiz Vechiato

1 Contextualização

As preocupações relacionadas ao projeto de ambientes informacionais digitais da *Web* contribuíram para o desenvolvimento de estudos realizados por comunidades de prática e pela comunidade científica de áreas do conhecimento que lidam com a interação humano-computador, como a Ciência da Computação e o Design de Interação, bem como com a interação dos sujeitos com a informação, como é o caso da Ciência da Informação (CI).

Dentre esses estudos, pode ser citada a Arquitetura da Informação (AI), que no contexto da *Web* e a partir de uma perspectiva interdisciplinar teve como uma das principais publicações a obra de Louis Rosenfeld e Peter Morville, em 1998, intitulada "*Information architecture for the World Wide Web*" (ROSENFELD; MORVILLE, 1998), que vem contribuindo, ao longo das suas quatro edições[4], com aportes metodológicos para o projeto de ambientes informacionais digitais.

Considerando que os estudos de AI partiram de uma comunidade de prática, é importante mencionar o desenvolvimento desses estudos pela comunidade científica, pois possibilitou avanços teóricos e metodológicos significativos ao longo dos anos, redimensionando seu campo de aplicação para ambientes não mais apenas digitais, mas também analógicos e híbridos, na perspectiva holística das ecologias informacionais complexas[5]. Isso trouxe

4 A edição mais atual desta publicação se intitula "Information architecture: for the web and beyond" e inclui o coautor Jorge Arango (ROSENFELD; MORVILLE; ARANGO, 2015). Por meio do título, é possível compreender que a publicação evoluiu considerando as estruturas de informação que não se apresentam apenas disponíveis na Web atualmente.

5 Segundo Oliveira e Vidotti (2016, p. 97), "[...] as Ecologias Informacionais Complexas são um conjunto de espaços e ambientes (analógicos, digitais ou híbridos), tecnologias (analógicas, digitais ou híbridas) e sujeitos, todos interligados e entrelaçados de maneira holística pela informação."

como consequência o desenvolvimento de quatro abordagens para a AI, sendo a Arquitetura da Informação Pervasiva a que está diretamente relacionada à evolução mencionada. (OLIVEIRA; VIDOTTI; BENTES, 2015).

Diante do desenvolvimento da AI, assistimos à gênese do conceito de *findability*, proposto por Peter Morville na obra *"Ambient findability"*, de 2005. (MORVILLE, 2005). Sua preocupação em mencionada obra foi evidenciar a dificuldade que muitos sujeitos possuem em encontrar informações que estão dispersas nos mais variados ambientes e dispositivos. Para o autor, a utilização de métodos e técnicas de AI é a solução para esta problemática.

Ainda na obra de Morville (2005), é possível identificar a ênfase nos processos informacionais ligados à interação dos sujeitos com a informação, mesmo que por um viés mais técnico que científico. Esse foi o motivo de estudar o conceito de *findability* e defender a tese (VECHIATO, 2013) de que referido conceito deveria ser incorporado aos estudos do campo da CI como um processo informacional dentro do fluxo de informação, com ênfase nos sujeitos informacionais.

Essa premissa subsidiou a proposta do termo "encontrabilidade da informação" (EI), por incorporar o *findability* ao campo da CI, embasado por aportes epistemológicos e teóricos que o sustentam no referido campo científico. Ademais, foi possível propor um conceito operatório para o termo "encontrabilidade da informação", reformulado em publicação posterior. Desse modo, entendemos que a EI "[...] sustenta-se fundamentalmente entre as funcionalidades de um ambiente informacional e as características dos sujeitos informacionais" (VECHIATO; VIDOTTI, 2014, p. 164).

Com vistas à compreensão dos fatores que influenciam a EI em ambientes informacionais, foram propostos treze atributos que estão relacionados aos aportes teóricos que sustentam o estudo na CI, os quais são apresentados no Quadro 1 que segue:

Quadro 1 - Atributos de Encontrabilidade da Informação

(continua)

Atributo	Descrição
Taxonomias Navegacionais	Utilizadas em estruturas de organização *top-down* (ambientes com características estáticas), referem-se à organização das categorias informacionais com vistas a facilitar a navegação e a descoberta de informações. Essas categorias, por exemplo, são organizadas geralmente em menus ou no corpo das páginas *Web*, nas comunidades e coleções de repositórios ou nas legendas utilizadas para descrição dos assuntos nas estantes das bibliotecas, organizadas previamente a partir de um sistema de classificação.
Instrumentos de controles terminológicos	Compreendem os vocabulários controlados, como os tesauros e as ontologias, para apoiar a representação dos recursos informacionais.
Folksonomias	Estão relacionadas à organização social da informação e propiciam ao sujeito a classificação de recursos informacionais, bem como encontrar a informação por meio da navegação (uma nuvem de *tags*, por exemplo) ou dos mecanismos de busca, ampliando as possibilidades de acesso. São utilizadas em estruturas de organização *bottom-up* (ambientes colaborativos). Quando associadas aos vocabulários controlados e às tecnologias semânticas, potencializam as possibilidades de encontrabilidade da informação.
Metadados	Compreendem a representação dos recursos informacionais e são armazenados em banco de dados para fins de recuperação da informação.
Mediação dos informáticos	Está associada ao desenvolvimento de sistemas, dispositivos, bancos de dados e interfaces com utilização de linguagens computacionais, com vistas à gestão e à recuperação da informação.
Mediação dos profissionais da informação	Ocorre em ambientes informacionais em que há sujeitos institucionais envolvidos na seleção, estruturação e disseminação da informação.
Mediação dos sujeitos informacionais	Está relacionada às ações infocomunicacionais que os sujeitos informacionais empreendem em quaisquer sistemas e ambientes informacionais, por exemplo, no que diz respeito à produção e à organização da informação e do conhecimento em ambientes colaborativos, gerados a partir de seus conhecimentos, comportamento e competências que caracterizam sua Intencionalidade.

Quadro 1 - Atributos de Encontrabilidade da Informação

(conclusão)

Affordances	Funcionam como incentivos e pistas que os objetos possuem e proporcionam aos sujeitos a realização de determinadas ações na interface do ambiente. Essas ações estão relacionadas a orientação, localização, encontrabilidade, acesso, descoberta de informações entre outras.
Wayfinding	Associado a orientação espacial, utiliza-se de aspectos que facilitam a localização, a encontrabilidade e a descoberta de informações por meio da navegação na interface do ambiente.
Descoberta de informações	Está condicionada aos demais atributos de encontrabilidade da informação no que diz respeito às facilidades que a interface (navegação e/ou mecanismos de busca) oferece para encontrar a informação adequada às necessidades informacionais do sujeito, bem como a possíveis necessidades informacionais de segundo plano.
Acessibilidade e Usabilidade	Relacionados à capacidade do sistema permitir o acesso equitativo à informação (acessibilidade) no âmbito do público-alvo estabelecido em um projeto com facilidades inerentes ao uso da interface (usabilidade).
Intencionalidade	A teoria da Intencionalidade fundamenta a importância em se enfatizar as experiências e habilidades dos sujeitos informacionais no projeto de ambientes e sistemas de informação.
Mobilidade, Convergência e Ubiquidade	Estão associados ao meio ambiente, externo aos sistemas e ambientes informacionais, mas que os incluem, dinamizando-os e potencializando as possibilidades dos sujeitos em encontrar a informação por meio de diferentes dispositivos e em diferentes contextos e situações.

Fonte: Vechiato, Oliveira e Vidotti (2016, p. 7).

Todos os atributos mencionados se alinham para o objetivo da encontrabilidade da informação pelo sujeito que utiliza um ambiente informacional.

Vale mencionar que a perspectiva da EI foi inovadora na tese de Vechiato (2013), embora já houvesse outros estudos sobre *findability*, ainda eram em uma perspectiva mais técnica que científica. Os estudos começaram a ser aplicados posteriormente a partir da elaboração de um *checklist* desenvolvido por Vechiato, Oliveira e Vidotti (2016), que viabilizou o desenvolvimento dos seguintes trabalhos:

EI em *web sites* (RODAS *et al.*, 2016; SILVA; VECHIATO; VIDOTTI, 2017); EI em repositórios institucionais (CUSTÓDIO; VECHIATO, 2016; CUSTÓDIO; VECHIATO, 2017; SANCHEZ; VECHIATO, 2017; FERNANDES; VECHIATO, 2017); EI em repositórios de dados (SANCHEZ; VIDOTTI; VECHIATO, 2017); EI na ecologia informacional complexa da Câmara dos Deputados (BRANDT; VECHIATO; VIDOTTI, 2018). Houve também avanços nas teorias associadas à EI, como: estudo específico a respeito do atributo *affordance* (TRINDADE; VECHIATO, 2016) e estudo específico a respeito do atributo *wayfinding* (MIRANDA; VECHIATO, 2017).

Considerando a contextualização realizada, este capítulo se destina a explorar as perspectivas de investigação da EI com ênfase no atributo Descoberta de Informação, que se apresenta incipientemente desenvolvido. Essas perspectivas dialogam com os Sistemas de Recomendação e com a Serendipidade, conforme será demonstrado no decorrer do capítulo.

2 Descoberta de informação como atributo da encontrabilidade da informação: perspectivas atuais

De forma autônoma, a EI ocorre na interação de um sujeito informacional com um mecanismo de busca e por meio da navegação / exploração em um sistema ou ambiente informacional digital, analógico ou híbrido no âmbito de uma ecologia informacional complexa.

Os atributos de EI estão associados aos processos informacionais e se misturam entre as características e funcionalidades do sistema e as características dos usuários, conforme demonstra o conceito apresentado na seção anterior.

A ideia central presente neste estudo é a de que, mesmo que um sistema funcione na perspectiva dos projetistas, desenvolvedores e gestores, nem sempre os sujeitos que farão uso desses sistemas conseguem encontrar facilmente a informação disponível, pois suas características não são consideradas no decorrer do processo de desenvolvimento.

Em um primeiro momento, a aplicação de testes de usabilidade pode auxiliar a identificar problemas na interação dos sujeitos com a interface dos

sistemas, mas ainda não são suficientes, pois não consideram aspectos que antecedem o uso do sistema em relação à experiência e ao comportamento informacional desses sujeitos.

Wilson (2000, p. 1, tradução nossa) define comportamento informacional como "[...] a totalidade do comportamento humano em relação às fontes e aos canais de informação, incluindo a busca ativa e passiva de informação e seu uso". Case (2007, p. 5, tradução nossa) entende que o comportamento informacional "[...] engloba a busca de informação e a totalidade de outros comportamentos não intencionais ou passivos (tais como "[descobrir] informação"), bem como comportamentos intencionais que não envolvem busca, por exemplo, evitar a informação".

Quando se alinha o conceito de Comportamento Informacional à EI, entende-se que, para que os sujeitos encontrem a informação que possibilitará satisfazer suas necessidades informacionais, podem ser ativados diferentes comportamentos em relação à informação e que o sucesso do processo dependerá das experiências, das atitudes e das habilidades dos sujeitos.

Nesse sentido, destaca-se o atributo Descoberta de Informação, compreendido, em um primeiro momento, como um comportamento não intencional ou passivo que propicia ao sujeito encontrar uma informação que é de seu interesse, mas não se caracteriza como a necessidade informacional que estava em primeiro plano, quando iniciou a interação com o sistema.

O embasamento teórico para a inclusão deste atributo nos estudos de EI se dá a partir do conceito de *information encountering* abordado por Erdelez (2006). A autora introduziu este conceito a partir de um estudo, em 1997, relacionado à aquisição acidental de informação (que compreendemos aqui como descoberta de informação), considerando que as pessoas podem encontrar informação sem uma intenção prévia (VECHIATO, 2013).

A autora define o termo *information encountering* "[...] como uma experiência memorável de descoberta inesperada de informações úteis ou interessantes." (ERDELEZ, 2006, p. 179, tradução nossa) e afirma que os ambientes da *Web*, por apresentarem uma ampla quantidade e variedade de informações, bem como facilidade de manipulação, proporcionam muitas vantagens para a descoberta de informação nessa perspectiva.

A partir deste conceito, Erdelez (2006) elaborou um modelo que pressupõe:

a) uma tarefa intencional de busca de informação relacionada a um problema que atua em primeiro plano para o sujeito (*foreground problem*);
b) um ou mais interesses, problemas e tarefas que atuam em um segundo plano para o sujeito informacional (*background interest/problem/task*), os quais não estão ativos no seu comportamento informacional.

Desse modo, em uma determinada situação, o sujeito informacional é submetido a uma interrupção de sua atenção, sendo transferido do seu *foreground problem*, ou seja, da busca intencional, para o *background problem*, ou seja, para a descoberta.

Erdelez acrescenta que o modelo envolve uma combinação de processos cognitivos, afetivos e comportamentais nos seguintes momentos: verificação da relevância da informação para o problema do segundo plano; interrupção do processo de busca original para examinar a informação localizada "acidentalmente" ou descoberta; armazenamento da informação, caso haja relevância; e retorno ao processo inicial de busca (ERDELEZ, 2006).

A autora ainda considera que interrupções e *looping* podem ocorrer, pois o processo varia de pessoa para pessoa. No ambiente *Web*, essas duas situações ocorrem muitas vezes, visto que, em um ambiente informacional, um sujeito pode não se lembrar de uma tarefa inicial, bem como pode alternar entre níveis de prioridade na busca de informações a todo o momento, isto é, uma busca intencional pode resultar na descoberta de informação que não estava atrelada à necessidade de informação inicial (ERDELEZ, 2006).

A descoberta de informação nesse sentido pode trazer sentimentos positivos para o sujeito, alterando seu comportamento, o que pode propiciar sua satisfação na interação com o ambiente. Desse modo, a compreensão desses comportamentos indica uma grande contribuição aos projetos de AI, fornecendo informações facilmente encontráveis por meio de busca intencional ou não intencional.

Alguns atributos de EI mencionados no Quadro 1 já estão diretamente relacionados a descoberta de informação, como:

a) taxonomias navegacionais: a estruturação lógica e a categorização adequada de assuntos podem auxiliar na descoberta de novas informações, desde que pensadas em conjunto com uma amostra de sujeitos potenciais e por meio de técnicas que explorem modelos mentais como o *card sorting*;
b) instrumentos de controle terminológico, metadados e *folksonomias*: a representação efetiva por metadados e o uso de vocabulários controlados, ontologias ou *folksonomias* auxiliam na encontrabilidade da informação por meio dos mecanismos de busca. Todavia podem ser utilizados para gerar recursos de navegação ou aprimoramento de busca que subsidiem a descoberta de recursos informacionais relacionados;
c) mediação dos informáticos e *affordances*: estão associadas ao desenvolvimento de recursos na interface baseados no perfil dos sujeitos, como os sistemas de recomendação, que serão aprofundados mais adiante. Esses recursos são apresentados com o apoio das *affordances*, que dão pistas de descoberta de novas informações aos sujeitos;
d) intencionalidade: refere-se às experiências dos sujeitos e à direcionalidade para a informação. O mapeamento do perfil dos sujeitos auxilia no desenvolvimento de recursos pelos desenvolvedores. Em ambientes colaborativos, a *folksonomia* é resultante da Intencionalidade dos sujeitos e contando apenas com sua própria mediação é possível que outros sujeitos descubram novas informações por meio da busca ou nuvem de *tags*.

Diante das relações apresentadas, é possível traçar perspectivas de investigação acerca do atributo Descoberta de Informação.

3 Sistema de recomendação e serendipidade: novas perspectivas para o atributo descoberta de informação

Os sistemas de recomendação facilitam a descoberta de informações no momento em que os sujeitos informacionais interagem com os mecanismos de busca ou navegação nas interfaces de sistemas e ambientes de informação, oferecendo a eles indicações de recursos informacionais relacionados e/ou personalizados durante esta interação.

Em sua tese de doutorado, Alvarez (2017, p. 127-128, grifo nosso) alinha a aplicação dos sistemas de recomendação à EI, quando apresenta seu conceito:

> Sistemas de Recomendação são aplicações concebidas para ambientes de informação digital, com capacidade para processar grandes volumes de dados, que atuam sugerindo recursos informacionais aos usuários, em contextos onde as alternativas de escolha são múltiplas e variadas, o que torna os processos de assimilação e processamento das informações, muito mais complexos. [...] **um Sistema de Recomendação está voltado a apoiar aos usuários no processo de encontrabilidade da informação, facilitando-lhes uma via alternativa para a localização de forma acidental ou não daquela informação que pode vir a satisfazer suas necessidades ou interesses informacionais conhecidos, declarados ou não, a partir da descoberta dos recursos informacionais armazenados nas bases de dados**, independente do ambiente, entorno ou contexto informacional onde o usuário esteja (tradicional, digital ou híbrido) e assim promover o acesso de forma rápida e continua à informação.

Os ambientes de *e-commerce* são potenciais na utilização desse recurso, propiciando aos sujeitos a descoberta de produtos relacionados ao que foi selecionado. Para isso, os dados do usuário são coletados (e aqui podemos alinhar ao atributo Intencionalidade) de duas formas, de acordo com Bisset Alvarez *et al*. (2016):

a) explícita: por meio da customização, o sujeito indica suas preferências e, desse modo, o sistema provê informação selecionada;
b) implícita: por meio do rastreamento das ações do usuário, estando ele cadastrado e suas ações coletadas por meio de perfil restrito, ou mesmo em ambientes abertos, por meio de seu registro de interação.

A aplicação dos sistemas de recomendação em ambientes informacionais pode ocorrer a partir de diferentes opções de filtragem de recursos, estratégias e abordagens (BISSET ALVAREZ *et al*., 2016) que são desenvolvidas pelos profissionais de informática que se utilizam da Intencionalidade dos sujeitos para coletar dados a respeito do seu perfil de interação, bem como de *affordances* para dar pistas aos sujeitos dos recursos relacionados, proporcionando a descoberta de informação.

O conceito de serendipidade, por sua vez, está diretamente associado a descoberta. Consolaro (2008, p. 24) descreve a respeito deste conceito:

> Muitas descobertas são atribuídas ao "acaso" ou "por acidente". Algumas foram mesmo por acaso, mas muitas outras assim consideradas são agora atribuídas à serendipidade. Quando um fenômeno, um evento, uma observação são detectados por mentes preparadas e adequadamente interpretadas podem gerar descobertas, inventos, teorias, enfim conhecimento para que o homem se adeque cada vez mais ao seu mundo. Quantas maçãs e outros frutos não caíram sobre a cabeça de pessoas ou de animais antes de Newton formular sua teoria sobre a lei da gravidade! Mas a sua mente preparada permitiu uma interpretação adequada do evento e sua transformação em conhecimento. Isto se chama "serendipidade". Pasteur afirmava: "No campo da observação, o acaso favorece apenas a mente preparada".

É possível evidenciar que muitas descobertas científicas foram impulsionadas pela serendipidade. Porém, a partir do relato do autor, o acaso é favorecido quando a mente está preparada para interpretar a ação que desencadeará a descoberta.

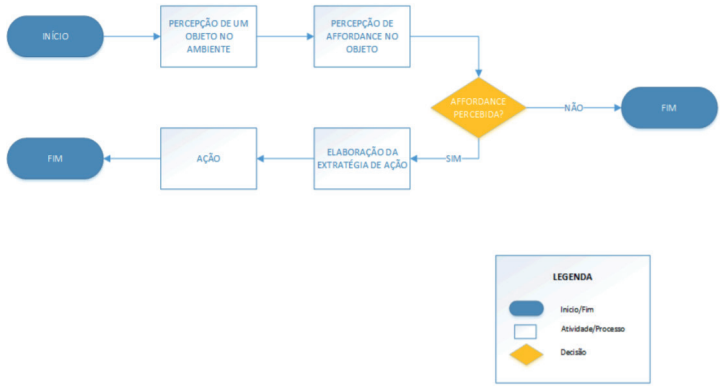
A partir dessa ideia é possível relacionar o conceito de serendipidade à noção de *affordance,* o referido conceito está associado ao reconhecimento de pistas presentes nos objetos e nos ambientes pelos sujeitos. O psicólogo americano

James Gibson elaborou a Teoria das *Affordances* a partir de seus estudos no campo da percepção visual. Para Gibson (*apud* OLIVEIRA, 2005, p. 90-91),

> [...] affordances são possibilidades de ação que o ambiente [ou objeto] oferece ao agente. Apesar de parecer uma definição simples, há características do conceito envolvidas nesta definição, que necessitam ser destacadas. Uma delas está baseada na ideia de que durante a interação com o ambiente, o agente percebe as possibilidades de ação e não as qualidades do ambiente. Além disso, a captação de tais possibilidades depende da escala corporal e das capacidades de ação do agente.

Para que uma *affordance* possa ser percebida, é necessário que o agente/sujeito identifique as ações possíveis em um determinado objeto ou ambiente. Do mesmo modo, para que haja serendipidade, é preciso ter a mente preparada para que ocorra a descoberta, e isso pode ocorrer em um primeiro momento a partir de *affordances*. O fluxograma exposto na Figura 1 apresenta o processo de percepção de *affordances*:

Figura 1 - Fluxo do processo de percepção de uma *affordance* e ação



Fonte: Trindade e Vechiato (2016, p. 907)

Percebe-se na figura que, caso o agente não consiga perceber a *affordance* em um dado objeto, ele não incidirá a ação correta a ponto de finalizar a ação com eficiência, podendo recorrer a estratégias de tentativa e erro, porém estas não configuram que a *affordance* foi percebida ou será percebida após várias tentativas.

No contexto da descoberta de informação, é extremamente importante que os sujeitos consigam perceber as *affordances* presentes. Isso auxiliará na serendipidade, porém não é o único aspecto que irá subsidiá-la.

De acordo com Makri *et al*. (2015), a serendipidade em ambientes informacionais digitais diz respeito à descoberta de informação que percebemos ser úteis, potencialmente úteis e inesperadas, e isso pode ocorrer quando estamos buscando por alguma informação, mas nos deparamos com outras, o que nos remete ao modelo "*information encountering*" de Erdelez (2006). Entretanto, para o autor, a descoberta também pode ocorrer quando não estamos buscando por nenhuma informação, mas sim apenas interagindo com o ambiente, o que sugere uma abordagem diferente da apresentada por Erdelez (2006), por não evidenciar uma necessidade informacional de primeiro plano.

Nesse momento, podemos relacionar a serendipidade à Intencionalidade. De acordo com Miranda (2010, p. 155):

> A Intencionalidade, segundo Sokolowski (200459), é a doutrina nuclear da fenomenologia, ela explica que cada ato de consciência que nós realizamos, cada ato de consciência que nós temos como algo intencional, é consciência de ou experiência de algo ou outrem.

Ao alinhar a teoria da Intencionalidade com a *findability* a partir de seus estudos, afirma que:

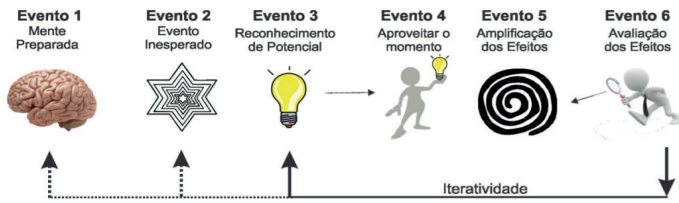
> A intencionalidade de informação para a findability significa direcionalidade de informação e se funda na experiência de cada sujeito [...]; a informação que é produzida é sempre acerca de e dirigida a, isto é, um sujeito, com a sua experiência, cria informação acerca de e dirigida a para atingir seus objetivos.

> E é nesse sentido, da capacidade da experiência do usuário, da consciência Intencional, que se baseia a web da inovação, do paradigma atual (MIRANDA, 2010, p. 273).

Mesmo que o foco da autora não seja a descoberta da informação, é possível compreender que a experiência do sujeito direciona suas ações no sistema no momento da busca intencional, não intencional ou na prática da serendipidade. Isso porque as experiências dos sujeitos possibilitarão que tenham ou não a mente preparada para a descoberta.

Paiva (2016) apresenta o modelo perceptivo da serendipidade que nos ajuda a compreender o processo de descoberta, demonstrado na Figura 2.

Figura 2 – Modelo perceptivo de serendipidade



Fonte: Paiva (2016, p. 38), adaptada de Lawley (2013).

Por meio desse modelo, percebemos que é necessário ter a mente preparada para o processo de descoberta. Desse modo, reforçamos o que já temos defendido em nossas pesquisas a respeito da investigação de aspectos que permeiam a Intencionalidade dos sujeitos informacionais.

Isso requer estudos que identifiquem traços da experiência do usuário e do comportamento informacional, de modo que a Arquitetura da Informação do ambiente não apenas satisfaça as necessidades informacionais, entendidas como aquelas que os usuários possuem no momento de uma busca intencional, mas também as derivadas de comportamentos não intencionais, que estimulem a serendipidade e, consequentemente, a descoberta de informações.

Diante do que foi abordado, um dos recursos importantes para que ocorra esse estímulo é o investimento em sistemas de recomendação que possibilitem

aos sujeitos obterem informações personalizadas e/ou relacionadas, baseadas na sua interação.

Para fins de sistematização conceitual, portanto, entendemos que a descoberta de informação é a recepção passiva e não intencional de informação pelo sujeito informacional, que proporcionará satisfação a ele, bem como a construção de conhecimento, visto que se trata de informação de seu interesse, obtida a partir da serendipidade, e que pode ser potencializada tanto por tecnologias e recursos que permitam mapear explícita ou implicitamente suas ações de interação quanto por estudos com sujeitos pertencentes ao público-alvo durante todo o processo de projeto do ambiente informacional.

4 Considerações finais

Este capítulo apresentou um conjunto de ideias que surgiram com o desenvolvimento de pesquisas em Encontrabilidade da Informação. Notávamos que o atributo Descoberta de Informação, embora extremamente relevante, estava incipientemente desenvolvido.

Desse modo, tecemos aqui perspectivas de investigação com ênfase neste atributo, buscando contribuir para seu aprimoramento. É claro que neste texto apenas damos um ponto de partida para a investigação, sendo necessário buscar mais aportes teóricos, especialmente a respeito da serendipidade, para consolidá-la no campo da Ciência da Informação.

A partir destas reflexões, este capítulo traz as seguintes perspectivas de investigação para o atributo Descoberta de Informação no contexto da Encontrabilidade da Informação:

a) os sistemas de recomendação devem ser entendidos como um recurso relevante a ser implementado em qualquer ambiente informacional digital, tendo em vista a ampla oportunidade de se constituir como um recurso de descoberta de informação. Ele se apoia na Intencionalidade dos sujeitos a partir de tecnologias que mapeiam sua interação, bem como em *affordances* que possibilitam ao sujeito interpretar as recomendações que podem estar alinhadas aos seus interesses; e

b) a serendipidade passa a ser constituir como um atributo de encontrabilidade da informação na perspectiva do sujeito informacional, visto que contribui para a descoberta de informação, mas não se caracteriza como ela propriamente dita, pois a descoberta em si ocorre no momento de interação com a interface. A serendipidade também depende do atributo Intencionalidade, pois indica o preparo da mente do sujeito para interpretar um recurso a ser descoberto. Outros atributos estão também associados a ela, como a usabilidade e a acessibilidade do ambiente, que promovem facilidades de interação; o *wayfinding*, que promove facilidades de orientação no ambiente; e as *affordances*, que indicarão as pistas que, se identificadas, poderão culminar na descoberta de informação.

Referências

BISSET ALVAREZ, E. *et al*. Os sistemas de recomendação, arquitetura da informação e a encontrabilidade da informação. **Transinformação**, Campinas, v. 28, n. 3, p. 275-286, set./dez. 2016. Disponível em: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0103-37862016000300275&lng=en&nrm=iso&tlng=pt. Acesso em: 15 jun. 2018.

BISSET, E. **Sistemas de recomendação para bibliotecas universitárias**: um aporte teórico da arquitetura da informação. 2017. Tese (Doutorado em Ciência da Informação) – Universidade Estadual Paulista "Júlio de Mesquita Filho", Marília, 2017. Disponível em: https://repositorio.unesp.br/handle/11449/151685. Acesso em: 15 jun. 2018.

BRANDT, M. B.; VECHIATO, F. L.; VIDOTTI, S. A. B. G. Encontrabilidade da Informação na Câmara dos Deputados. **Em Questão**, v. 24, p. 41-64, 2018. Disponível em: <http://seer.ufrgs.br/index.php/EmQuestao/article/view/71734>. Acesso em: 15 jun. 2018.

CASE, D. O. **Looking for information.** 2. ed. Amsterdam: Elsevier; Academic Press, 2007.

CONSOLARO, A. Adaptação ao mundo, avalanche de informações e a serendipidade na Odontologia. **R Dental Press Ortodon Ortop Facial**, Maringá, v. 13, n. 3, p. 23-27, maio/jun. 2008. Disponível em: <http://www.scielo.br/pdf/dpress/v13n3/a03v13n3.pdf>. Acesso em: 15 jun. 2018.

CUSTODIO, N. C.; VECHIATO, F. L. Encontrabilidade da informação em repositórios institucionais: uma proposta de instrumento de avaliação. **Revista Informação na Sociedade Contemporânea**, Natal, v. 1, p. 1-17, 2017. Disponível em: <https://periodicos.ufrn.br/informacao/article/view/12284>. Acesso em: 15 jun. 2018.

CUSTODIO, N. C.; VECHIATO, F. L. Mediação infocomunicacional no contexto da encontrabilidade da informação: uma análise do processo de autoarquivamento no repositório institucional da Universidade Federal do Rio Grande do Norte. **Biblionline,** João Pessoa, v. 12, p. 3-13, 2016. Disponível em: <http://periodicos.ufpb.br/ojs/index.php/biblio/article/view/28235>. Acesso em: 15 jun. 2018.

ERDELEZ, S. Information encountering. In: FISHER, K. E.; ERDELEZ, S.; McKECHNIE, L. (ed.). **Theories of information behavior.** Medford: Information Today Inc, 2006. p. 179-184.

FERNANDES, W. M.; VECHIATO, F. L. Encontrabilidade da informação no Repositório Institucional da Unesp: uma avaliação com dispositivos móveis. *In*: Encontro Nacional de Pesquisa em Ciência da Informação, 18., 2017, Marília. **Anais...** p. 1-8. Disponível em: <http://enancib.marilia.unesp.br/index.php/xviiienancib/ENANCIB/paper/viewFile/406/1271>. Acesso em: 15 jun. 2018.

LAWLEY, J. Maximising serendipity: the art of recognising and fostering unexpected potential – a systemic approach to change. *In*: **Neuro Linguistic Psychotherapy and Counselling Association 2013**. [S.l.: s.n.], 2013.

MAKRI, S. et al. Observing serendipity in digital information environments. *In*: ASIS&T ANNUAL MEETING, 78., 2015, St. Louis. **Proceedings** [...]. St. Louis: ASIST, 2015. p. 1-10

MIRANDA, E. B. S.; VECHIATO, F. L. Wayfinding em ambientes informacionais: proposta de recomendações para avaliação no contexto da encontrabilidade da informação. **Revista Informação na Sociedade Contemporânea**, Natal, v. 1, p. 1-21, 2017. Disponível em: https://periodicos.ufrn.br/informacao/article/view/10976. Acesso em: 15 jun. 2018.

MIRANDA, M. K. F. de O. **O acesso à informação no paradigma pós-custodial**: da aplicação da intencionalidade para findability. 2010. Tese (Doutorado em Informação e Comunicação em Plataformas Digitais) – Faculdade de Letras, Universidade do Porto, Porto, 2010.

MORVILLE, P. **Ambient findability.** Sebastopol: O'Really, 2005.

OLIVEIRA, F. I. da S. **Affordances:** a relação entre agente e ambiente. 2005. 99 f. Dissertação (Mestrado em Filosofia) – Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2005.

OLIVEIRA, H. P. C.; VIDOTTI, S. A. B. G. Dos ambientes informacionais às ecologias informacionais complexas. **Informação & Sociedade**, João Pessoa, v. 26, n. 1, p. 91-101, jan./abr. 2016. Disponível em: <http://www.periodicos.ufpb.br/ojs/index.php/ies/article/view/29438/16509>. Acesso em: 15 jun. 2018.

OLIVEIRA, H. P. C.; VIDOTTI, S. A. B. G.; BENTES, V. **Arquitetura da informação pervasiva.** São Paulo: Cultura Acadêmica, 2015.

PAIVA, F. A. P. de. **Estudo do conceito de serendipidade como base para novas abordagens ao problema da convergência prematura.** 2016. 110f. Tese (Doutorado em Engenharia Elétrica e de Computação)

– Centro de Tecnologia, Universidade Federal do Rio Grande do Norte, Natal, 2016.

RODAS, C. M. et al. Encontrabilidade da informação: uma análise a partir da tecnologia de eye-tracking. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 17., 2016, Salvador. **Anais** [...]. Salvador: UFBA, 2016. p. 1-20.

ROSENFELD, L.; MORVILLE, P. **Information architecture for the world wide web.** Sebastopol: O'Really, 1998.

ROSENFELD, L; MORVILLE, P.; ARANGO, J. **Information architecture**: for the web and beyond. 4. ed. Sebastopol: O'Really, 2015.

SANCHEZ, F. A.; VECHIATO, F. L. Encontrabilidade da informação em repositórios digitais: um enfoque nos repositórios institucionais da USP, Unesp e UNICAMP. *In*: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 18., 2017, Marília. **Anais** [...]. Marília: Unesp, 2017. p. 1-8. Disponível em: http://enancib.marilia.unesp.br/index.php/xviiienancib/ENANCIB/paper/viewFile/358/1270. Acesso em: 15 jun. 2018.

SANCHEZ, F. A.; VIDOTTI, S. A. B. G.; VECHIATO, F. L. Contribuições da arquitetura da informação e da encontrabilidade da informação para o projeto e a avaliação de repositórios de dados. *In*: ENCONTRO IBÉRICO EDICIC, 8., 2017, Coimbra. **Anais** [...] Disponível em: http://sci.uc.pt/eventos/atas/edicic2017.pdf. Acesso em: 15 jun. 2018.

SILVA, N. B. P.; VECHIATO, Fernando Luiz; VIDOTTI, Silvana Aparecida Borsetti Gregorio. Encontrabilidade da Informação no ambiente Technology, Entertainment and Design (TED). **Informação & Informação** (Online), v. 22, p. 481-496, 2017.

TRINDADE, A. S. C. E.; VECHIATO, F. L. Da percepção a ação: affordances como elementos facilitadores para a encontrabilidade

da informação em bibliotecas. In: SEMINÁRIO EM CIÊNCIA DA INFORMAÇÃO, 6. 2016, Londrina. **Anais** [...]. Londrina: UEL, 2016. p. 903-915. Disponível em: http://www.uel.br/eventos/cinf/index.php/secin2016/secin2016/paper/view/309. Acesso em: 15 jun. 2018.

VECHIATO, F. L. **Encontrabilidade da informação**: contributo para uma conceituação no campo da ciência da informação. 2013. 206 f. Tese (Doutorado em Ciência da Informação) – Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2013. Disponível em: <https://www.marilia.unesp.br/Home/Pos-Graduacao/CienciadaInformacao/Dissertacoes/Tese_de_Doutorado_-_Fernando_Luiz_Vechiato.pdf>. Acesso em: 15 jun. 2018.

VECHIATO, F. L.; VIDOTTI, S. A. B. G. **Encontrabilidade da informação**. São Paulo: Cultura Acadêmica, 2014. (Coleção PROPG Digital UNESP). Disponível em: http://www.culturaacademica.com.br/catalogo/encontrabilidade-da-informacao/. Acesso em: 16 jun. 2018.

VECHIATO, F. L.; OLIVEIRA, H. P. C.; VIDOTTI, S. A. B. G. Arquitetura da informação pervasiva e encontrabilidade da informação: instrumento para a avaliação de ambientes informacionais híbridos. *In*: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 17., 2016, Salvador. **Anais** [...]. Salvador: UFBA, 2016.

WILSON, T. D. Human information behaviour. **Informing science**, v. 3, n. 2, p. 49-55, 2000.

Análise, tendências e perspectivas da produção científica em curadoria digital

Sandra de Albuquerque Siebra
Vildeane da Rocha Borba

1 Introdução

O mundo cada vez mais digital dos dias atuais vem modificando as formas como as pessoas trabalham, comunicam-se, divertem-se, acessam e usam a informação. De fato, o perfil dos interagentes (mais ativo, exigente e produtor e consumidor de informações), as tecnologias de rede e as facilidades tecnológicas oferecidas no contexto da *world wide web (www)* e a produção intensiva de dados de pesquisa e de dados abertos no contexto do *e-goverment* têm resultado na geração de enormes quantidades de dados[6] em todas as áreas do conhecimento, englobando os contextos acadêmico, científico, profissional, governamental e pessoal.

Além disso, essas quantidades estão aumentando significativamente. Em previsão feita por Gantz e Reinsel (2012), o universo digital vai crescer em um fator de 300, de 2005 a 2020, indo de 130 exabytes para 40.000 exabytes, o que equivale a dizer que o universo digital duplicará a cada dois anos, até 2020.

Essa profusão de dados nato-digitais sendo produzidos, assim como de objetos analógicos sendo trazidos para o universo digital por meio da digitalização, tem sido a promessa de acesso onipresente/ubíquo. Porém a fragilidade inerente ao universo digital, a evanescência de mídias e arquivos (que podem ser corrompidos), a rápida obsolescência de software e hardware, a necessidade de sistemas de arquivos e metadados bem construídos, além dos meandros dos direitos de propriedade intelectual, colocam todos esses materiais em risco e oferecem pouca esperança de longevidade para informações que não são intencionalmente preservadas (OLIVER; HARVEY, 2016). Assim, a herança cultural, a memória

6 Dados podem ser definidos como qualquer informação em formato binário. Essa definição é intencionalmente ampla e extensível, além de ter uma estreita relação com o produto de pesquisas científicas. Essa definição inclui objetos digitais e bases de dados (DCC, 2004-2015).

individual e coletiva, o conhecimento científico e o governo aberto dependem da preservação de dados e objetos digitais confiáveis e autênticos.

Nesse cenário, tornam-se relevantes as pesquisas, discussões e trabalhos no contexto da curadoria digital, que tem sido concisamente definida como o gerenciamento ativo e a preservação de material digital para garantir acesso e uso a longo prazo (ABBOTT, 2008). De fato, observa-se o crescimento da curadoria digital como campo de pesquisa nos últimos anos. No entanto, há poucos trabalhos direcionados ao estudo de mudanças no campo e para identificar tendências predominantes e lacunas na literatura de pesquisa (SANTOS, 2014; KIM, 2015).

Dessa forma, buscando contribuir com o preenchimento dessa lacuna, o objetivo desta pesquisa foi analisar a produção científica nacional e internacional sobre curadoria digital, no período de 2004 a 2017, nas bases de dados Base de Dados Referencial de Artigos de Periódicos em Ciência da Informação (BRAPCI) e Scopus.

2 Curadoria digital

O termo curadoria digital foi usado pela primeira vez em 2001 (KIM, 2015), no seminário "Digital Curation: Digital Archives, Libraries and e-Science", como um termo para destacar algumas das mudanças que eram necessárias em abordagens para a curadoria de artefatos digitais, em oposição a artefatos analógicos.

A curadoria digital é considerada um termo polissêmico que abrange diversas atividades de uma variedade de profissões, instituições e setores (LEE; TIBBO, 2011). Inclusive, muitas pessoas admitem usar os termos preservação digital, curadoria digital e até mesmo o arquivamento e gestão de dados digitais de forma intercambiável (LAZORCHAK, 2011).

No entanto, a curadoria digital é um conceito mais inclusivo do que o arquivamento digital ou a preservação digital. Ela aborda uma ampla gama de processos aplicados aos objetos digitais ao longo de sua vida útil, começando antes da criação dos objetos digitais, definindo padrões para o planejamento da coleta de dados que resulta em objetos digitais "preparados para curadoria", que estão na melhor condição possível para garantir que possam ser mantidos e usados no

futuro (OLIVER; HARVEY, 2016). Por exemplo, se são objetos digitalizados, eles estarão obedecendo a critérios e padrões adequados de resolução, tonalidade de bits, formato de arquivo, entre outros.

Por isso, a curadoria digital é definida como o desenvolvimento de um conjunto de técnicas que abordam "o que vem antes da preservação e o que vem depois - ou seja, como os dados são criados e como são usados (...) e como serão usados e por quem, no futuro" (HARVEY, 2010, p. 16). Essencialmente, a curadoria digital é um processo holístico que analisa todo o ciclo de vida dos dados, desde a criação até a manutenção e na preparação de como os dados podem ser ajustados ou migrados no futuro. Trata-se de tomar decisões informadas desde o início para garantir dados úteis mais tarde.

Além disso, a curadoria digital enfatiza a adição de valor a conjuntos de dados e objetos digitais, pela inserção de metadados ou anotações adicionais, para que possam ser mais bem recuperados e reutilizados. De fato, a curadoria digital está preocupada com e é aplicável a uma ampla gama de objetos digitais. Ela aplica-se igualmente tanto a objetos digitais complexos, vinculados a outros recursos em diversos formatos, grandes conjuntos de dados científicos ou conjuntos de dados que estão mudando a cada segundo, quanto a objetos digitais relativamente simples, como os documentos estáticos (ex: livros), geralmente manipulados por bibliotecas e arquivos. O que difere da maioria das práticas de arquivamento de dados e preservação digital que foram desenvolvidas para documentos estáticos e não são transferidas com sucesso para dados mais complexos.

Logo, nesta nova era de pesquisas orientadas por dados, novas estratégias e processos são necessários para lidar com a ampla gama de dados criados e mantidos por muitos tipos diferentes de comunidades de usuários. Levando tudo isso em conta, pode-se afirmar que

> A curadoria digital está preocupada com o gerenciamento ativo de dados, enquanto continuar a ser de interesse acadêmico, científico, de pesquisa, administrativo e / ou pessoal, com os objetivos de dar suporte a reprodutibilidade, reutilização e agregação de valor a esses dados, gerenciá-los desde o seu

> ponto de criação até que seja determinado que ele não será mais útil, e garantir sua acessibilidade, preservação, autenticidade e integridade ao longo do tempo. (OLIVER; HARVEY, 2016, p. 8).

Yakel (2007) argumenta que a preservação digital é um subconjunto da curadoria digital. Para Foscarini et al. (2010, p.1), a preservação digital visa assegurar que os objetos digitais "[...] possam ser significativamente reproduzidos ao longo do tempo, apesar da evolução das representações, mecanismos, do rápido avanço das tecnologias e das continuamente emergentes expectativas dos usuários". A curadoria digital é holística e abrange todo o ciclo de vida do conteúdo digital, sendo a preservação um ponto específico desse ciclo de vida (CAPLAN, 2011), o que pode ser observado nos modelos que vêm sendo utilizados no contexto da curadoria digital (ex: *DCC Lifecycle Model* (HIGGINS, 2008)).

3 Metodologia

Esta pesquisa teve cunho quali-quantitativo e descritivo, utilizando como técnica de análise de dados a bibliometria e a análise documental (MICHEL, 2009). A coleta de dados sobre a produção científica em "Curadoria Digital" foi realizada na Scopus e na BRAPCI, em junho de 2018. A escolha da Scopus deveu-se a abrangência mais elevada de quantidade de documentos e da BRAPCI como uma opção para entender o panorama brasileiro de produções na temática em questão.

O termo de busca utilizado em ambas as bases foi "Curadoria Digital", em língua portuguesa na BRAPCI e em língua inglesa na Scopus. Na Scopus, o campo de busca utilizado abrangeu o título, resumo e palavras-chave, na BRAPCI foram utilizados todos os campos.

Assim, foram recuperados 325 documentos na Scopus, porém, restringindo para a pesquisa dos tipos documentais artigos de periódicos e trabalhos apresentados em eventos, abrangendo o período de 2004 a 2017, totalizou-se 244 trabalhos, dos quais 149 artigos e 95 trabalhos apresentados em eventos. Na BRAPCI foram localizados 26 documentos, abrangendo o período de 2012 a 2017, dos quais 6 são trabalhos apresentados em eventos e 20 são trabalhos

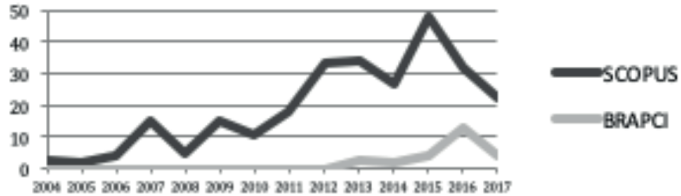
publicados em artigos de periódicos. O período analisado foi composto a partir do ano que as bases apresentaram registros sobre a temática, até 2017. O ano de 2018 foi excluído por ainda estar em andamento. Neste sentido, a abrangência pesquisada foi de 2004 a 2017.

Os dados da Scopus foram processados pelo software *Vantage point*, versão 9.0, e os da BRAPCI manualmente, analisando os seguintes dados: evolução do número de publicações/ano em Curadoria digital; as principais fontes que publicaram sobre Curadoria Digital; país dos autores com o maior número de publicações; afiliação dos autores com o maior número de publicações; trabalhos mais citados e temáticas relacionadas com a Curadoria Digital. Adicionalmente, foi realizada a análise documental do título, veículo de publicação, palavras-chave e resumo dos artigos recuperados, a fim de ter um panorama sobre as abordagens dadas em âmbito nacional e internacional e as áreas que colaboraram com as pesquisas em curadoria.

4 Resultados e discussões

Apesar do termo curadoria digital ter surgido em 2001 (KIM, 2015), a distribuição temporal da produção sobre Curadoria Digital na Scopus apresenta registros a partir de 2004, revelando o maior número de publicações em 2015 e o menor em 2005 (Gráfico 1). Em 2004, duas das três publicações dizem respeito ao *Digital Curation Centre* (DDC), que é um Centro de expertise em Curadoria digital que foi fundado justamente em março de 2004, no Reino Unido.

Gráfico 1 – Evolução do número de publicações por ano em curadoria digital na Scopus (n=244) e BRAPCI (n=26)

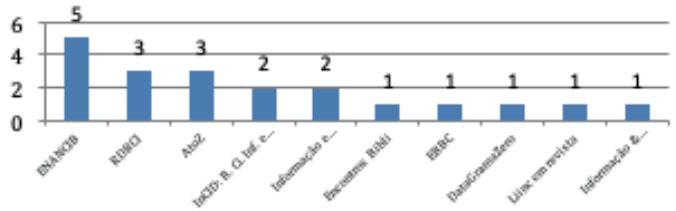


Fonte: Autoria própria (2018).

Na BRAPCI, os registros mostram publicações a partir de 2013[7], o que indica que as pesquisas sobre curadoria digital começaram a ser realizadas mais tardiamente no Brasil, tendo o maior número de publicações ocorrido em 2016.

O Gráfico 2 apresenta as dez fontes com maior quantidade de publicações na temática no Brasil. Foram encontradas 16 fontes, das quais 2 delas são eventos e 14 são periódicos. Os dois eventos com publicações sobre Curadoria digital são Encontro Nacional de Pesquisa em Ciência da Informação (ENANCIB) e Encontro Brasileiro de Bibliometria e Cientometria (EBBC), enfatizando que o ENANCIB obteve o maior número de publicações, o que é condizente com a cobrança que tem sido feita aos pesquisadores pelos órgãos de fomento de terem publicações em eventos reconhecidos pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) e periódicos, de preferência, qualificados. Dos 14 periódicos brasileiros, 13 deles são vinculados a área da Ciência da Informação e apenas a Revista P2P & Inovação é vinculada ao Grupo de Pesquisa Economias Colaborativas e Produção P2P no Brasil, do Instituto Brasileiro de Informação em Ciência e Tecnologia – IBICT.

Gráfico 2 – As 10 fontes com maior número de publicações na pesquisa do termo curadoria digital na BRAPCI

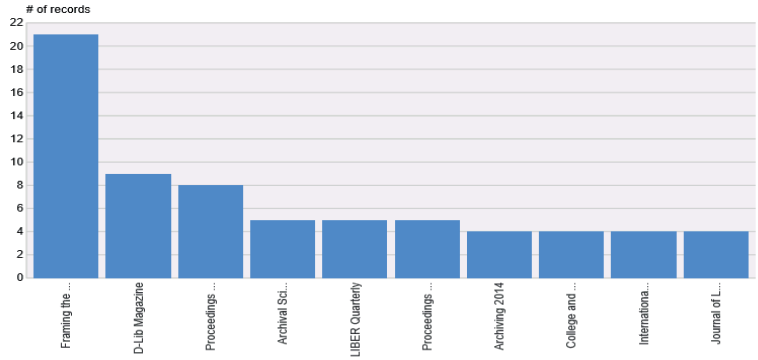


Fonte: Autoria própria (2018).

7 Enfatiza-se que existe um artigo de 2012 não registrado na base em questão, de autores brasileiros, cujo título é "Curadoria digital: um novo patamar para preservação de dados digitais de pesquisa" de autoria de Sayão e Sales, relacionando os estudos de curadoria digital ao de dados de pesquisa e, consequentemente, às necessidades de pesquisa e aprofundamento de estudos na temática no Brasil.

Na Scopus foram constatadas 146 diferentes fontes com publicações sobre curadoria digital, das quais 58 foram publicadas em diferentes eventos e 88 em diferentes periódicos (Gráfico 3). A Revista com o maior número de publicações sobre a temática foi o *D-Lib Magazine* com 9 publicações e o evento *Framing the Digital Curation Curriculum Conference* com 21 artigos.

Gráfico 3 – As 10 fontes com maior número de publicações na pesquisa do termo curadoria digital na Scopus

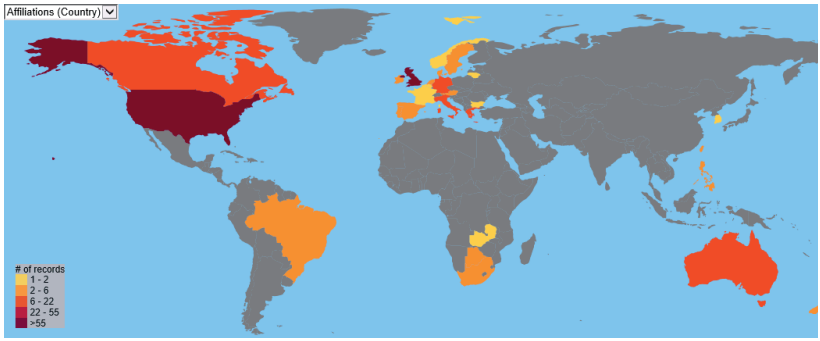


Fonte: Autoria própria (2018).

As publicações sobre Curadoria Digital têm sido mais realizadas nos Estados Unidos (Figura 1), sendo esse o país com maior quantidade de publicações sobre a temática (109 títulos). O Reino Unido está em segundo lugar, com 55 títulos indexados pela Scopus, principalmente pelo desenvolvimento do *Digital Curation Centre*. O Brasil está no décimo terceiro lugar, com 3 publicações nesta base.

Na totalidade, as produções advindas da Scopus foram escritas por 414 autores, afiliados a 254 instituições, sendo as duas primeiras instituições afiliadas a *School of Information and Library Science*, da Universidade da Carolina do Norte, e o *Digital Curation Centre* (DCC), da Universidade de Glasgow e de Edinburgh, no Reino Unido.

Figura 1 – Locais com publicações do termo curadoria digital na Scopus



Fonte: Autoria própria (2018).

Na BRAPCI, as produções contemplam os escritos de 44 autores, cujas principais afiliações se vinculam a Universidade Federal do Estado do Rio de Janeiro (UFRJ), a Universidade Federal de Pernambuco (UFPE) e a Universidade Estadual Paulista Júlio de Mesquita Filho (UNESP). Os 2 principais autores com o maior número de publicações sobre a temática são SALES, L. F. e SAYÃO, L. F., com 4 publicações indexadas na BRAPCI. Já na Scopus, os 2 principais autores que publicam sobre a temática são LEE, C. A e TIBBO, H. R., com 13 e 12 trabalhos, respectivamente, ambos afiliados a *School of Information and Library Science*, da Universidade da Carolina do Norte. Porém os autores com o maior número de citações (Quadro 1) são BRYAN HEIDORN, P. e TRANT, J., o primeiro vinculado a Universidade do Arizona nos Estados Unidos e o segundo vinculado ao *Archives and Museum Informatics*, também localizado nos Estados Unidos.

Entre os termos mais frequentes na produção pesquisada se destacam os termos Curadoria digital, Preservação Digital, Bibliotecas Digitais, Armazenamento Digital, Curadoria, Gerenciamento da Informação, e-learning, Curadoria de Dados e Humanidades Digitais.

Quadro 1 – Os 5 trabalhos com o maior número de citações na Scopus

AUTOR	TÍTULO	FONTE	ANO	CITAÇÃO
BRYAN HEIDORN, P.	The emerging role of libraries in data curation and e-science	Journal of Library Administration	2011	44
TRANT, J.	Emerging convergence? Thoughts on museums, archives, libraries, and professional training	Museum Management and Curatorship	2009	34
HEDGES, M., BLANKE, T., HASAN, A.	Rule-based curation and preservation of data: A data grid approach using iRODS	Future Generation Computer Systems	2009	26
TENOPIR, C., SANDUSKY, R.J., ALLARD, S., BIRCH, B.	Academic librarians and research data services: Preparation and attitudes	IFLA Journal	2013	25
KERNE, A., WEBB, A.M., SMITH, S.M., MOELLER, J., DAMARAJU, S.	Using metrics of curation to evaluate information-based ideation	ACM Transactions on Computer-Human Interaction	2014	21

Fonte: Autoria própria (2018).

O termo Curadoria Digital é o mais frequente, justicado pela temática pesquisada. O termo preservação digital possui relação direta com a curadoria digital, fazendo parte dos processos e ciclos de curadoria digital, como já havia sido afirmado por Yakel (2007). Além disso, frequentemente, a preservação digital faz parte da descrição do contexto histórico que antecede a curadoria digital.

Quadro 2 – As 10 palavras-chave mais frequentes da produção sobre curadoria digital na Scopus e BRAPCI

(continua)

PALAVRAS-CHAVE SCOPUS	FREQUÊNCIA	PALAVRAS-CHAVE BRAPCI	FREQUÊNCIA
Digital Curation	157	Curadoria Digital	13
Digital Preservation	42	Ciência da Informação	11
Digital Libraries	37	Preservação Digital	7
Digital Storage	28	E-Science	6

Quadro 2 – As 10 palavras-chave mais frequentes da produção sobre curadoria digital na Scopus e BRAPCI

(conclusão)

Curation	22	Tecnologia da Informação e Comunicação	3
Information Management	22	Biblioteconomia	3
E-Learning	16	Dados de Pesquisa	3
Data Curation	14	Ciência Aberta	3
Digital Humanities	14	Design da Informação	2
Education	14	Curadoria de Conteúdo	2

Fonte: Autoria própria (2018).

Os termos Curadoria e Curadoria de Dados são terminologias encontradas para o debate da temática. Importante enfatizar que o termo Curadoria é bastante utilizado na área da Ciência da Informação (CI), em especial para atividades de curadoria em exposição e eventos na Museologia, e a curadoria de dados está vinculada, especialmente no Brasil, a dados de pesquisa e a *e-science*. Já no exterior a curadoria de dados, na produção científica analisada, engloba, além dos dados de pesquisa, também os dados abertos governamentais, ainda poucos explorados no contexto da CI no Brasil.

Os termos Bibliotecas Digitais, Gerenciamento de Informação e Armazenamento Digital possuem uma alta frequência no contexto analisado. O primeiro devido ao desenvolvimento de acervos em meio digital, influenciado pelo movimento de acesso aberto e pela ciência aberta e pelos desafios que esse novo cenário traz para os arquivos, bibliotecas e instituições memoriais. Os demais termos são tratados no sentido dos desafios e estratégias que podem ser utilizadas para gerir e armazenar (ambos os termos partem dos processos de curadoria digital) grandes volumes de dados produzidos e disseminados. Em diversos artigos, estudos de caso envolvendo essas etapas do processo de curadoria são apresentados.

Os termos *e-learning* e Educação foram bastante utilizados, porque especialmente, mas não exclusivamente, nos anos de 2014 e 2015 boa parte dos artigos trabalhava formas de preparar a força de trabalho para curadoria digital. Nos artigos são discutidos o perfil e as competências necessárias ao curador digital; as mudanças necessárias em currículos de graduação; a criação de cursos de mestrado e doutorado em curadoria digital e são apresentados projetos de cursos, treinamentos e conferências para discutir questões educacionais e de implementação da curadoria na prática. Nesse período, destacam-se as publicações do Carolina Digital Curation Curriculum Project (projeto DigCcurr) da Universidade da Carolina do Norte (DIGCCURR PROJECT).

O termo Tecnologia da Informação e Comunicação se relaciona com o uso da tecnologia nos processos de curadoria digital e o Design da Informação com a melhor forma de trabalhar o conteúdo curado para ser apresentado aos interagentes da informação. Ressalta-se que, apesar de os artigos da SCOPUS apresentarem diversas pesquisas aplicadas, nas quais são contruídas ferramentas, frameworks e protótipos tecnológicos, esse termo é uma palavra-chave pouco utilizada nos artigos, talvez porque os pesquisadores considerem que o uso da tecnologia é inerente ao contexto digital.

Por fim, o termo humanidades digitais é uma terminologia contemporânea, constituído como uma comunidade de prática, no domínio das Ciências Humanas e Sociais que defende o acesso livre de dados e seus metadados, formatos e resultados de pesquisa, utilizando e migrando para os meios digitais o trabalho tradicional e seus usos potenciais, por isso tem relação próxima com as propostas da curadoria digital.

Adicionalmente, com a realização da análise documental dos resumos da produção científica sobre curadoria digital, por meio dos registros contidos na BRAPCI, percebe-se que no Brasil, como observado por Siebra, Borba e Miranda (2016), há três abordagens mais utilizadas para curadoria: a curadoria de dados de pesquisa (que atende ao contexto da e-science), a curadoria de conteúdo ou informação (que tem seu foco na mediação, a fim de amenizar o cenário de explosão informacional) e a curadoria digital com foco na preservação para acesso

e uso futuro. Ressalta-se que a maioria dos artigos ainda são teóricos e poucas experiências e estudos de caso foram encontrados.

Já na Scopus, o panorama dos artigos mostra que nos primeiros anos (2004-2010) a maioria das publicações abordava a criação de instituições (com destaque para o DCC) e projetos (com destaque para o DigCcurr) para trabalhar as questões relacionadas a curadoria. Nesse período surgem os primeiros modelos e propostas de políticas e padrões.

No período de 2011 a 2013, os artigos abordaram o papel das bibliotecas, arquivos e instituições na curadoria, o desenvolvimento de guias, tutoriais e materiais de referência, assim como iniciou-se a discussão sobre questões educacionais e de competências, que foi intensificada, como já citado, nos anos de 2014 e 2015. Até esses anos, as publicações tinham a maioria cunho teórico ou teórico-prático e é possível ver uma mistura de ênfase em questões técnicas, legais, sociais, comerciais e comportamentais.

Com o cenário mais bem estabelecido, as publicações dos anos de 2016 e 2017 tiveram um enfoque mais prático e aplicado, tratando, em sua maioria, do desenvolvimento e uso de ferramentas e protótipos; aplicação de frameworks e modelos; realização de estudos de caso e relatos de experiência, o que mostra um amadurecimento da área. Um fato curioso é que poucos artigos abordam a curadoria de dados de pesquisa ou a curadoria de conteúdo, especificamente. Também na análise documental verificou-se que as áreas que mais trabalharam em conjunto com CI na área de curadoria foram a Computação, a Educação, as Ciências Médicas, a Comunicação, a Sociologia e a Arqueologia.

5 Considerações finais

A curadoria de objetos digitais, sejam culturais, memoriais, educacionais, científicos ou governamentais, é um dos desafios centrais que vem desde o início do século XXI (TIBBO, 2012; HARLEY, 2010). Porém é inegável que as duas últimas décadas testemunharam o desenvolvimento de várias agendas de pesquisa e um amplo progresso em direção a arquiteturas robustas de repositórios (ex: EPrints; DSpace; iRODS), bem como o desenvolvimento de ferramentas, modelos, padrões

e estratégias para preservação (ex: DCMI, METS, PREMIS, DCC Lifecycle Model) e progresso rumo à curadoria digital confiável e sustentável.

Além disso, algumas instituições (ex: DCC e DigCcurr) vêm colaborando com o desenvolvimento de bases sólidas para pesquisa e evolução contínua. Porém, para o estabelecimento da área, especialmente no Brasil, ainda há necessidade de pesquisa rigorosa, da realização de muito mais experimentação, além da disseminação de melhores práticas e informações sobre padrões. Também urge a capacitação e preparação de um crescente grupo de profissionais para curadoria digital, requerendo, em alguns momentos, o repensar de currículos e cursos.

Adicionalmente, é preciso enfatizar que, quase sempre, a curadoria digital é, na prática, um esforço colaborativo e que os projetos de pesquisa de curadoria geralmente precisam envolver uma ampla gama de indivíduos com várias habilidades, conhecimentos e perspectivas. Ademais, a tecnologia deve ser ação fluindo de teorias e princípios, pois, naturalmente, o termo curadoria combina as facetas técnica (digital) e conceitual (curadoria). Logo, o desenvolvimento da tecnologia deve estar sempre a serviço do objetivo maior de preservar o conteúdo digital para usos e usuários futuros. A tecnologia deve ser vista como o meio, não como o objetivo em si, pois, no contexto da curadoria, embora muitas funções tenham aspectos técnicos, nenhuma delas pode ser realizada apenas pela tecnologia, muitas exigem extensa intervenção humana e julgamento, pelo menos por enquanto.

Referências

ABBOTT, D. "What is Digital Curation?". *In*: **DCC Briefing Papers**: introduction to curation. Edinburgh: Digital Curation Centre, 2008. Disponível em: http://www.dcc.ac.uk/resources/briefing-papers/introduction-curation/what-digital-curation. Acesso em: 25 jun. 2018.

BRYAN HEIDORN, P. The emerging role of libraries in data curation and e-science. **Journal of Library Administration**, v.51, n.7-8, p. 662-672, 2011. DOI: 10.1080/01930826.2011.601269

CAPLAN, P. **Semantics**: digital preservation vs. digital curation. 2011. Disponível em: https://groups.google.com/forum/?fromgroups= - topic/digital-curation/ehppkZT9XGs. Acesso em: 25 jun. 2018.

DIGCCURR PROJECT. **Preserving access to our digital future**: building an international digital curation curriculum. Disponível em: https://ils.unc.edu/digccurr/. Acesso em: 25 jun. 2018.

DIGITAL CURATION CENTRE (DCC). **The DCC Curation Lifecycle Model**. Edinburgh: Digital Curation Centre, 2004-2015. Disponível em: http://www.dcc.ac.uk/resources/curation-lifecycle-model/. Acesso em: 25 jun. 2018.

FOSCARINI, F. *et al*. On the notion of Genre in digital preservation. *In*: AUTOMATION IN DIGITAL PRESERVATION, Dagstuhl, 2010. **Proceedings** [...]. Dagstuhl: [s.n.], 2010.

GANTZ, J.; REINSEL, D. The digital universe in 2020: big data, nigger digital shadows, and biggest growth in the far east. **IDC**, New York, v. 3, dec. 2012. Disponível em: https://www.emc.com/collateral/analyst-reports/idc-digital-universe-united-states.pdf. Acesso em: 25 jun. 2018.

HARVEY, R. **Digital curation**: a how-to-do-it manual. New York: NealSchuman Publishers, 2010.

HEDGES, M.; BLANKE, T.; HASAN, A. Rule-based curation and preservation of data: A data grid approach using iRODS. **Future Generation Computer Systems**, v. 25, n. 4, p. 446-452, 2009.

HIGGINS, S. Digital curation: the emergence of a new discipline. **The International Journal of Digital Curation**, v. 6, n. 2, 2008. Disponível em: http://www.ijdc.net/index.php/ijdc/article/ view/184. Acesso em: 10 Abr. 2018.

KERNE, A.; WEBB, A. M.; SMITH, S. M.; MOELLER, J.; DAMARAJU, S. Using Metrics of Curation to Evaluate Information-Based Ideation.

ACM Transactions on Computer-Human Interaction, v. 21, n. 3; p. 1-48, 2014. DOI: 10.1145/2591677.

KIM, J. Growth and trends in digital curation research: the case of the international journal of Digital Curation. **ASIST**, Seattle, v. 51, n. 1, 2015. Disponível em: https://onlinelibrary.wiley.com/doi/full/10.1002/meet.2014.14505101074. Acesso em: 25 jun. 2018.

LAZORCHAK, B. **Digital preservation, digital curation, digital stewardship**: what's in (some) names? the signal, 2011. Disponível em: http://blogs.loc.gov/digitalpreservation/2011/08/digital-preservation-digital-curation-digital-stewardship-what%E2%80%99s-in-some-names/. Acesso em: 25 jun. 2018.

LEE, C. A.; TIBBO, H. Where's the archivist in digital curation?: exploring the possibilities through a matrix of knowledge and skills. **Archivaria**, New York, v. 72, p. 123-168, 2011.

MICHEL, M. H. **Metodologia e pesquisa científica em ciências sociais**: um guia prático para acompanhamento da disciplina e elaboração de trabalhos monográficos. São Paulo: Atlas, 2009.

OLIVER, G; HARVEY, D. R. **Digital curation**. Chicago: ALA Neal-Schuman, 2016.

SANTOS, T. N. C. **Curadoria digital**: o conceito no período de 2000 a 2013. 2014. Dissertação (Mestrado em Ciência da Informação) - Universidade de Brasília, Brasília, 2014. Disponível em: http://repositorio.unb.br/bitstream/10482/17324/1/2014_ThayseNataliaCantanhedeSantos.pdf. Acesso em: 25 jun. 2018.

SIEBRA, S. A.; BORBA, V. R.; MIRANDA, M. K. F. O. Curadoria digital: um termo interdisciplinar. **Informação & Tecnologia**, v. 3, p. 21-38, 2016.

TENOPIR, C.; SANDUSKY, R. J.; ALLARD, S.; BIRCH, B. Academic librarians and research data services: Preparation and attitudes. **IFLA Journal**, v. 39, n. 1, p. 70-78, 2013. DOI: 10.1177/0340035212473089

TIBBO, H. R. Placing the horse before the cart: conceptual and technical dimensions of digital curation. **Historical Social Research**, v. 37, n. 3, p. 187-200, 2012.

TRANT, J. Emerging convergence? Thoughts on museums, archives, libraries, and professional training. **Museum Management and Curatorship**, v. 24, n.4, p. 369-387, 2009. DOI: 10.1080/09647770903314738

YAKEL, E. Digital curation. **OCLC Systems and Services**: International digital library perspectives, v. 23, n. 4, p. 335-340, 2007.

Gestão da qualidade de dados conectados: fundamentos e metodologia

Leonardo Castro Botega
José Eduardo Santarem Segundo
Jéssica Oliveira de Souza Ferreira Melo
Caio Saraiva Coneglian

1 Introdução

Bizer, Heath e Berners-Lee (2009) descrevem dados conectados como a prática de utilizar a Web para criar links entre dados de diferentes fontes utilizando tecnologias e conceitos da Web Semântica. No entanto, os conjuntos de dados publicados podem variar na qualidade, podendo conter estruturas precisas, ou seja, dados bem formatados e livres de erros, ou conjuntos dotados de diversos problemas de qualidade, como incompletude, imprecisão, links quebrados etc. Neste contexto, qualidade pode ser considerada um conjunto de requisitos necessários para um dado atender às exigências, de acordo com objetivos específicos.

Uma das medidas mais conhecidas para evitar problemas de qualidade nos dados conectados, independentemente do domínio dos dados, é descrita por Berners-Lee (2006), que define quatro regras que expõem expectativas de comportamento para publicação que, quando atendidas, promovem a interconexão dos dados. Caso contrário, limita a reutilização desses dados. As regras estabelecidas pelo autor são:

a) utilizar URI (Uniform Resource Identifier) para nomear recursos;
b) utilizar HTTP (HyperText Transfer Protocol) como URI, de modo que tais dados possam ser encontrados;
c) prover informações úteis utilizando os padrões RDF (Resource Description Framework), SPARQL (Protocol and RDF Query Language);
d) incluir links que guiem a outros recursos URIs, de modo que o usuário possa encontrar mais informações relacionadas.

Além das quatro regras definidas por Berners-Lee (2006), a literatura aponta problemas de qualidade não somente nos dados, mas também na estrutura provida para sua publicação, fator que pode dificultar seu acesso e até mesmo inviabilizar sua utilização, evidenciando o fato de que a qualidade consiste em um fator de extrema importância.

O World Wide Web Consortium (W3C) fornece abrangente conteúdo visando orientar o processo de construção das informações a serem publicadas em bases de dados conectados abertos (LOD), abreviação em inglês para *Linked Open Data*, a fim de evitar erros de qualidade como: formatos de dados errados, links quebrados, criação de URIs, guias para utilização de padrões de metadados, ontologias etc. Porém, após quase uma década da criação do LOD, ainda é possível encontrar conjuntos na rede de dados que apontam para links quebrados e problemas como os citados acima, alguns dos quais se propagam desde a criação do projeto.

No contexto de dados conectados, os problemas de qualidade que têm afetado os conjuntos de dados publicados podem ser considerados obstáculos não somente para sua utilização, mas também para aplicações que fazem uso de tais dados. Adicionalmente, definições quanto às dimensões de qualidade para os dados conectados e para os métodos de avaliação são disponibilizadas, porém indaga-se como tais problemas e dimensões acontecem e afetam os conjuntos de dados conectados.

É também um desafio para a área de Ciência da Informação (CI) determinar de que modo os problemas de qualidade contradizem os princípios de utilização dos dados conectados e qual é a proporção de tais problemas de acordo com categorias específica de conjuntos de dados.

Conforme cunhado por Borko (1968), investigar propriedades, comportamento, meios de processamento de informações para acessibilidade, usabilidade, armazenamento, recuperação, interpretação, transmissão, armazenamento e utilização são temas de grande interesse para a comunidade de CI. Assim, considerando que os dados conectados dotados de qualidade contribuem de forma relevante para diversos temas de interesse da CI, as aplicações computacionais inteligentes, a detecção de problemas de qualidade neste tipo de

dados, bem como meios de prevê-los, mensurá-los ou eliminá-los representam a principal motivação desta pesquisa.

O estudo justifica-se pela evidente presença de problemas de qualidade que afetam os conjuntos de dados e a necessidade de uma descrição específica de como tais problemas acontecem em dados considerados livres de problemas de qualidade publicados no diagrama LOD (MENDES; MÜHLEISEN; BIZER, 2012; BIZER; CYGANIAK, 2009; ZAVERI *et al*., 2013; KONTOKOSTAS *et al*., 2013; RULA; ZAVERI, 2014).

Desta maneira, este capítulo tem como objetivos: (1) aplicar uma metodologia de avaliação de qualidade de dados conectados no contexto de dados criminais da Grécia presente no LOD, como forma de provar e validar os conceitos apresentados e (2) incluir e demonstrar aspectos referentes à representação da qualidade de dados conectados, no mesmo contexto e modelo ontológico de dados criminais.

2 Aplicação da metodologia de avaliação de qualidade de dados conectados

Esta seção tem como objetivo apresentar o uso da metodologia de avaliação de qualidade de dados conectados de Souza (2017), bem como ilustrar a aplicação de cada dimensão de qualidade, suas métricas e o cálculo de atribuição dos índices de qualidade. Para conduzir a avaliação foi utilizado um conjunto de dados criado pela Polícia Nacional da Grécia, que descreve recursos sobre ocorrências criminais de diversas categorias. Visto que o conjunto possui um grande volume de dados, foi selecionado um conjunto controlado, composto de 100 triplas, as quais descrevem mais de 80 tipos de crimes diferentes.

2.1 *Interlinking*

Esta dimensão tem como objetivo avaliar problemas em links RDF. A Fórmula 1 é utilizada para obter o resultado do índice de *interlinking* a partir de duas métricas.

$$I = \frac{m_1 + m_2}{2} \tag{1}$$

2.1.1 Métrica 1: detectar links de boa qualidade

Essa avaliação é conduzida levando em consideração os princípios para identificar links de boa qualidade, sendo esses os que atendem aos seguintes requisitos: o primeiro consiste em não utilizar *namespaces* que fazem uso de conjuntos ou identificadores numéricos para os recursos. Quando utilizados *namespaces*, as URIs não podem ser referenciáveis, visto que apenas o domínio que as criou conseguiria identificá-las. Uma URI é referenciável quando é possível recuperar o recurso identificado por ela. O segundo consiste em não utilizar nomes dos *hosts*, extensão de páginas, detalhes sobre o desenvolvimento na URI que não apresentem informações sobre o recurso. E, por fim, utilizar chaves naturais nas URIs: utilizar identificadores significativos para o domínio do conjunto de dados, como a combinação do nome e sobrenome do recurso (HEATH; BIZER, 2011; SAUERMANN; CYGANIAK; VÖLKEL, 2007; BERNERS-LEE, 2008).

A Fórmula 2 apresenta o cálculo para o índice da métrica, onde *U* representa o total de URIs e *Up* a quantidade de URI presente com problema de qualidade.

$$m_1 = \frac{100 - 100}{100} = 0 \tag{2}$$

Todas as URIs analisadas cumpriram com apenas um dos três requisitos de utilização de links de boa qualidade, o que consequentemente caracteriza problema de qualidade na URI de todos os recursos, resultando em um índice de 0%. O resultado aplicado na fórmula proposta é apresentado na Fórmula 3.

$$m_1 = \left(\frac{(\sum U) - (\sum Up)}{\sum U} \right) * 100 \tag{3}$$

2.1.2 Métrica 2: detectar a existência de links para fontes externas

De acordo com a proposta dos dados conectados, o ideal é que os conjuntos de dados tenham pelo menos 50 links para fontes externas e essa métrica é disposta em duas etapas:

a) é realizada uma verificação na página do conjunto no *Datahub*, se tais informações não forem disponibilizadas;
b) uma busca no arquivo dos dados é realizada visando encontrar o conteúdo sobre propriedades utilizadas para apontar links externos.

Datahub é uma ferramenta de gerenciamento e publicação de grandes coleções de dados e todos os conjuntos inseridos no diagrama LOD possuem um registro no *Datahub*, e, quando inseridos no diagrama, subentende-se que foram disponibilizadas informações descritivas sobre eles.

$$m_2 = \left(\frac{(\sum Ln) - (\sum Lp)}{\sum Ln} \right) * 100 \quad (4)$$

O cálculo dessa métrica é apresentado na Fórmula 4, onde *Ln* consiste no total de links necessários, subtraída a quantidade de links presentes *Lp,* e então o resultado é transformado em porcentagem.

Propõe-se a aplicação da fórmula quando a quantidade de links presentes for menor do que a quantidade necessária, visto que, quando for maior, o requisito foi cumprido, totalizando os 50% correspondentes dessa métrica.

De modo geral, cada métrica corresponde a 50% do índice local dessa dimensão; a pontuação da métrica 1 é realizada para verificar a porcentagem de URIs sem problema de qualidade e a métrica 2 para identificar se o conjunto de dados possui ao menos 50 links. E então é efetuada a soma dos valores das duas métricas para chegar ao índice local de qualidade dessa dimensão. A avaliação de ambas as métricas resulta em um valor quantitativo.

O requisito para que um conjunto de dados seja considerado de qualidade consiste em cumprir o requisito de ter no mínimo 50 links para fontes externas, e o conjunto de dados tinha mais de 20 mil links externos.

$$I = \frac{0 - 100}{2} = 50 \tag{5}$$

Foi constatado que esse conjunto de dados possui links para 11 conjuntos diferentes, todos incluídos no diagrama LOD. Assim, esta métrica atinge o seu total no índice geral de *interlinking* de 50% e o índice geral desta dimensão foi de 50%, conforme apresentado na aplicação da Fórmula 5 para calcular o valor desta dimensão, apresentada a seguir.

2.2 Licenciamento

2.2.1 Métrica 1: detectar a existência de uma licença na documentação do conjunto de dados

Um requisito para os dados abertos e para que o conjunto de dados seja inserido no diagrama do projeto LOD é a disponibilização da licença na página do *Datahub*, independentemente do tipo, avaliada qualitativamente quanto à ausência ou à presença da licença.

Durante a verificação do conjunto de dados exemplo, constatou-se que foi disponibilizada a licença, uma vez que o campo da indicação descrevia licença do tipo *Creative Commons Attribution Share-Alike*, e a *tag licence-metadata* foi inserida no conjunto de *tags* do conjunto de dados.

2.2.2 Métrica 2: especificar a licença correta, se está atribuída sob a licença original

Considerando o fato de que existem tipos diferentes de licenças, essa métrica, também qualitativa, tem como objetivo que o avaliador distinga se a licença disponibilizada é de fato a licença que deveria ter sido disponibilizada, se é a correta, de acordo com o domínio do conjunto de dados.

Essa verificação também é realizada na página do conjunto de dados; ao clicar no link da licença disponibilizada, o avaliador é redirecionado para uma

página de descrição da licença utilizada. Desse modo, o avaliador poderá então distinguir se o que foi disponibilizado é a licença correta ou não.

A não disponibilização da licença invalida a aplicabilidade dessa métrica, porém, visto que é um requisito obrigatório, quando acontecer, é avaliada de modo negativo, ou seja, como se não fosse a licença correta disponibilizada para o tipo de dados.

Essa métrica tem como objetivo disponibilizar a licença para que o avaliador distinga se ela é a licença correta para os dados disponibilizados, dimensão definida com o objetivo de auxiliar o usuário final dos dados, a fim de definir se a licença condiz com a sua aplicação desejada. Visto que não representa o contexto desta avaliação, foi invalidada na análise do conjunto de dados criminais.

2.3 Consistência

2.3.1 Métrica 1: verificar se os dados estão de acordo com a especificação da ontologia

Esta métrica tem como objetivo identificar valores contraditórios, ou seja, inconsistentes com a classe ou propriedade que a especifica. Tal problema de qualidade acontece quando a propriedade descreve valores contraditórios com o que foi especificado na classe e/ou propriedade do vocabulário descrito.

A avaliação é conduzida sob o arquivo do conjunto de dados; assim, é fundamental para a análise que esse arquivo tenha sido disponibilizado na página do conjunto de dados no Datahub. Nesse sentido, é importante que o avaliador tenha um pouco de conhecimento dos termos utilizados para a descrição de conteúdo, a fim de listar quais propriedades são passíveis de gerar inconsistências e contradições quando utilizadas juntas. Após buscar pelos termos, o avaliador irá analisar se as classes e propriedades foram utilizadas de modo satisfatório.

Ao final, é atribuído um índice por meio do cálculo apresentado na Fórmula 6, onde *TTi* consiste na somatória de classes/propriedades da amostragem, da qual é subtraída a somatória de *Ti*, que representa a quantidade de itens inconsistentes. O resultado é então dividido pelo total de itens *TTi*, que ao final é multiplicado por 100 para encontrar a porcentagem final.

$$CONm_1 = \left(\frac{(\sum TTi) - (\sum Ti)}{\sum TTi} \right) * 100 \quad (6)$$

Conforme ressaltado no início do capítulo, a utilização de propriedades e classes OWL auxilia a aplicação da avaliação da dimensão. No caso da utilização de diferentes e/ou vocabulários próprios, a avaliação consiste em detectar propriedades e classes do vocabulário do conjunto de dados que se contradizem e não poderiam ser utilizadas na mesma declaração.

As informações sobre o vocabulário utilizado foram encontradas na página oficial do projeto LOD da polícia grega desenvolvido pelo Ministério da Ordem Pública e Proteção ao Cidadão[8]. Por meio da documentação foi constatado que o vocabulário possui apenas 4 classes: crimes (11.070 recursos), *crimetype* (41 recursos), *gad* (15 recursos) e *regions* (15 recursos).

Foi conduzida uma análise sobre cada termo do vocabulário, tanto nas propriedades quanto nas classes, a fim de verificar prováveis inconsistências em relação às classes e propriedades que foram de fato utilizadas no conjunto de dados. Tal inconsistência é definida como classes ou propriedades que descrevem recurso ou dado X, porém tem como valor um recurso ou dado *Y*, entretanto, não foi constatado nenhum tipo de problema de consistência quanto a métrica 1.

Desse modo, essa métrica atingiu 100% de seu índice de qualidade, porém, visto que representa metade do valor local, seu valor inteiro corresponde a 50%, conforme apresentado na aplicação da Fórmula 7, onde o resultado da soma das propriedades (19) e classes (4) é subtraído do total de propriedades e classes inconsistentes (0) e o resultado é então dividido pelo total de classes e propriedades e multiplicado por 100.

$$CONm_1 = \frac{(23) - (0)}{23} * 100 = 100 \quad (7)$$

8 Disponível em: http://greek-lod.math.auth.gr/police/. Acesso em: 8 jun. 2018.

2.3.2 Métrica 2: verificar o tipo de dados permitido para a propriedade

Cada campo permite um tipo diferente de dados; esta análise consiste em verificar se o tipo do dado inserido condiz com o tipo permitido para o campo.

Cada métrica representará 50% do índice local dessa dimensão, onde é efetuada a soma dos índices que é convertido para a porcentagem final de consistência. O cálculo dessa métrica é apresentado na Fórmula 8, onde *Tm* consiste no total de atributos que possuem um tipo pré-determinado de dados, *Tmi* representa o total de propriedades com problema de consistência.

$$CONm_2 = \left(\frac{(\sum Tm) - (\sum Tmi)}{\sum Tm} \right) * 100 \tag{8}$$

Das 19 propriedades existentes no vocabulário, apenas 4 definiam um tipo específico de dados. Ao analisar todos os 100 recursos da amostragem, pode-se afirmar que não houve problemas de consistência quando esta métrica.

Desse modo, essa métrica atingiu 100% de seu índice de qualidade, porém, visto que representa metade do valor local, seu valor inteiro corresponde a 50%, conforme apresentado na aplicação da Fórmula 9, onde o total de atributos com tipo pré-determinados é subtraído do total de atributos inconsistentes e o resultado é o resultado é então dividido pelo total de atributos pré-determinados e multiplicado por 100.

$$CONm_2 = \frac{(4) - (0)}{4} * 100 = 100 \tag{9}$$

2.4 Precisão sintática

2.4.1 Métrica 1: detectar o uso de regras sintáticas

Essa verificação é conduzida para analisar o tipo de caracteres ou modelo de valores permitidos. Cada domínio possui tipos específicos de dados que devem conter uma quantidade ou modelo exato, como, por exemplo, *Digital Object Identifier* (DOI), que consiste em um identificador único para identificar

entidades físicas, digitais ou abstratas e prover um link persistente para sua localização na internet.

O cálculo do índice dessa dimensão é apresentado na Fórmula 10, onde *Tam* consiste no total de atributos com tipos e formatos específicos de dados e *Tap* no total de valores com problemas de precisão; o resultado é convertido para valor final de precisão local.

$$PSm_1 = \left(\frac{(\sum Tam) - (\sum Tap)}{\sum Tam} \right) * 100 \quad (10)$$

Após uma verificação das propriedades do vocabulário, constatou-se a não existência de propriedades que têm um modelo específico, visto que a dimensão não pôde ser avaliada em razão da inexistência deste tipo de dados. Desse modo, nesse caso específico, tal dimensão foi excluída do índice geral de qualidade.

2.5 Precisão semântica

2.5.1 Métrica 1: detectar a utilização de propriedades inexistentes

Um problema possível de acontecer consiste na declaração de classes e propriedades não existentes no conjunto do vocabulário/linguagem utilizada para descrição dos dados, com exceção, ressalte-se, da criação de classes e propriedades específicas de um vocabulário criado exclusivamente para determinados conjuntos de dados, o que implicaria a disponibilização e especificação na web.

A análise foi realizada por meio de uma verificação de cada termo utilizado na amostragem, sendo que para cada termo uma busca nos termos da documentação foi realizada, tanto nas classes quanto nas propriedades e nenhum problema de precisão semântica foi detectado.

2.5.2 Métrica 2: uso incorreto das propriedades

Essa métrica tem como objetivo verificar a utilização inapropriada de atributos entre dois recursos (ou seja, como uma relação) e de relações

(propriedades) como valores literais. Para exemplificar, no caso da utilização do OWL, pode acontecer a utilização errada dos termos *owl:DatatypeProperty* e *owl:ObjectProperty*: *datatypeProperty* descreve uma propriedade (atributo) e *objectProperty* define uma propriedade de relacionamento entre dois recursos, ou seja, a relação.

Foi verificada cada propriedade e classe das triplas analisadas e constatou-se que não ocorreu este tipo de problema. Assim, todas as classes utilizadas eram de fato classes, bem como todas as propriedades utilizadas.

2.6 Completude

A primeira etapa da avaliação de completude é abordada nas métricas 1, 2, 3 e a segunda etapa é descrita na métrica 4. O índice local de completude é composto pelas cinco métricas propostas, conforme apresenta a Fórmula 11, cada uma equivalendo a 20% do valor total; assim, caso todas as métricas estejam completas, o índice local de completude é 100%. Esse 100% é então adaptado para o seu valor equivalente dentro de um intervalo de 0 a 25% do domínio geral, que é quanto cada uma das cinco dimensões ocupa.

$$Com = \frac{m_1 + m_2 + m_3 + m_4 + m_5}{5} \tag{11}$$

No caso de exceções inesperadas, como a não disponibilização ou não utilização dos dados necessários para conduzir a avaliação, a métrica é invalidada e não aplicada; assim, o valor total da métrica é dividido pela quantidade de cada métrica passível de aplicação. Caso aconteça de não ser disponibilizado o necessário em todas as métricas, o valor de completude do conjunto de dados é de 0%.

Os cálculos das métricas 1, 2 e 3 foram realizados do seguinte modo: primeiro foi realizada uma verificação em cada tripla das classes e propriedades utilizadas, excluindo as repetições. Constatou-se que o vocabulário possui 23 termos sendo 19 propriedades e 4 classes, destes foram utilizadas 2 classes e 5 propriedades.

2.6.1 Métrica 1: completude de esquema

Essa análise obtém como resultado o quão completo o esquema da ontologia está, em questão de propriedades e de classes utilizadas. Para realizar a verificação é necessário o arquivo da ontologia, que é disponibilizado na página do conjunto de dados. Finalmente, é então verificado quantas classes e propriedades foram representadas.

O cálculo é apresentado na Fórmula 12, onde *Cr* consiste na soma das classes e propriedades representadas na amostragem a ser analisada e *Tc* o total de classes e propriedades.

$$COMm_1 = \left(\frac{\sum Cr}{Tc} \right) * 100 \tag{12}$$

O total de classes e propriedades utilizadas foram 7 e o percentual foi de 29,16%, conforme apresentado na Fórmula 13.

$$COMm_1 = \left(\frac{7}{24} \right) * 100 = 29{,}16 \tag{13}$$

2.6.2 Métrica 2: completude de propriedade

Nessa métrica é obtida a quantidade de propriedades que foram utilizadas na ontologia. O cálculo é apresentado na Fórmula 14, onde Pr consiste no total de propriedades representadas e TP no total de propriedades

$$COMm_2 = \left(\frac{\sum Pr}{TP} \right) * 100 \tag{14}$$

Das 19 propriedades apenas 5 foram utilizadas e o percentual foi de 26,31%, conforme apresentado na Fórmula 15.

$$COMm_2 = \left(\frac{5}{19} \right) * 100 = 26{,}31 \tag{15}$$

2.6.3 Métrica 3: completude de população

Consiste em obter a completude das classes, ou seja, quantas foram utilizadas em relação ao total de classes disponíveis para utilização. O cálculo é apresentado na Fórmula 16, onde *Tcp* consiste no número de objetos do mundo real representados e *TC* no número total de objetos do mundo real.

$$COMm_3 = \left(\frac{\sum Tcp}{TC} \right) * 100 \tag{16}$$

Foram utilizadas 2 classes e o percentual foi de 50%, conforme a Fórmula 17.

$$COMm_3 = \left(\frac{2}{4} \right) * 100 = 50 \tag{17}$$

2.6.4 Métrica 4: quantidade de metadados necessários a serem disponibilizados no *Datahub*

A avaliação dessa métrica é realizada por meio da verificação das informações disponibilizadas no registro. O cálculo é apresentado na Fórmula 18, em que *Ip* consiste na quantidade de informações presentes, de onde são subtraídas as adicionais *Ia*; o resultado obtido é dividido pelo total de informações necessárias *In*, de onde são subtraídas as adicionais *Ia*.

$$COMm_4 = \left(\frac{(\sum Ip) - (Ia)}{(\sum \mathrm{In}) - Ia} \right) * 100 \tag{18}$$

Uma exceção acontece no caso das tags *lodcloud.nolinks*, *lodcloud.unconnected*, *lodcloud.needinfo*, *lodcloud.needsfixing*, que são obrigatórias, porém não são utilizadas quando o problema retratado por elas não for aplicável ao conjunto de dados em avaliação. Desse modo, tais propriedades não são incluídas no cálculo quando essa exceção acontece.

Dos 35 atributos, 4 foram excluídos do cálculo por não serem aplicáveis ao conjunto de dados em avaliação e 6 por serem atributos adicionais. Dos 25 restantes, 4 não foram informados. Tendo assim 25 atributos obrigatórios,

totalizando 95% de completude, conforme apresentado na Fórmula 19, onde as informações presentes (26) foram subtraídas das adicionais (6) e então divididas pela subtração total de informações obrigatórias (25) e informações adicionais (6).

$$COMm_4 = \left(\frac{(26)-(6)}{(25)-(6)} \right) * 100 = 96 \tag{19}$$

2.7 Avaliação temporal (*timeliness* e volatilidade)

2.7.1 Métrica 1: verificar quão atuais são os dados

Tem como objetivo obter duas informações temporais sobre o *conjunto de dados*, sendo que tais são classificadas em duas categorias:

a) *timeliness*: verifica se o dado é atual para o contexto no qual será utilizado;
b) volatilidade: caracteriza a frequência na qual os dados, no caso o conjunto de dados, variam no tempo (BOUZEGHOUB, 2004; BATINI; SCANNAPIECO, 2016).

A verificação quanto à *timeliness* é realizada do seguinte modo: a data da última atualização é subtraída da data atual, resultando na idade do conjunto de dados. A partir do resultado o avaliador poderá verificar se, de acordo com suas tarefas e prioridades, o dado poderá ser considerado atual ou não.

A volatilidade consiste em contabilizar a quantidade de vezes em que o conjunto de dados foi atualizado, quando o resultado é 0, significa que não foi atualizado nenhuma vez, o que, dependendo da sua data de criação, pode significar que o conjunto de dados estejam desatualizado. Consequentemente, quando o valor for maior que 0, significa que o conjunto de dados já foi atualizado.

A última atualização do conjunto de dados ocorreu no dia 30 de julho de 2016, e a avaliação foi realizada no dia 30 de abril de 2018, resultando então em 669 dias corridos. Em relação à volatilidade, foi disponibilizada uma única informação sobre atualização, deduzindo-se, assim, que o conjunto de dados foi atualizado uma vez em 30 de julho de 2016.

2.8 Índice global de qualidade

O índice global é composto pelo índice local das dimensões quantitativas: precisão sintática, completude, consistência e *interlinking*. Visto que o índice global é composto por quatro dimensões, cada uma é responsável por 25% do valor total. Desse modo, o índice local, que consiste em um valor de 0 a 100%, é reajustado de modo a prover seu valor equivalente dentro do limite, em que 25% corresponde a 100%.

O resultado do índice geral de qualidade do conjunto de dados avaliado foi de 66,62%, conforme apresentado na Fórmula 20. Para realizar o cálculo foram utilizados os índices das dimensões de consistência, completude (Fórmula 21) e interliking (Fórmula 22).

$$Ig = \frac{50 + 49{,}86 + 100}{3} = 66{,}62 \tag{20}$$

$$Com = \frac{29{,}16 + 26{,}31 + 50 + 95}{4} = 89{,}86 \tag{21}$$

$$I = \frac{0 - 100}{2} = 50 \tag{22}$$

Assim, o resultado do índice local de cada dimensão, sem a equivalência, foi o seguinte:

a) precisão sintática = invalidada;
b) consistência = 100%, visto que as métricas 1 e 2 atingiram seu valor total;
c) *interlinking* = 50%;
d) completude = 49,86%.

O Quadro 1 apresenta um cartão proposto para auxiliar o processo de avaliação de acordo com cada dimensão. QT corresponde à avaliação quantitativa e QL à qualitativa. Propõe-se a inserção dos índices de locais de qualidade, para as dimensões quantitativas, para o cálculo do índice de qualidade geral.

3 Representação da qualidade no contexto de dados conectados

Nesta seção, será abordada uma forma de representação de informações de qualidade em um modelo semântico, como estas estão em relação ao cenário de dados conectados, integração e representação das avaliações junto aos *datasets* considerados.

Quadro 1 – Índices de qualidade de cada dimensão

Dimensão	Métricas	Avaliação	Pontuação	Local
Interlinking	M1	QT	0%	50%
	M2	QT	50%	
Licenciamento	M1	QL	Não se aplica	X
	M2	QL	Não se aplica	
	M3	QL	Não se aplica	
Consistência	M1	QT	50%	100%
	M2	QT	50%	
Precisão sintática	M1	QT	Invalidada	Invalidada
Precisão semântica	M1	QL	Não se aplica	X
	M2	QL	Não se aplica	
Completude	M1	QT	29,16%	49,86%
	M2	QT	26,31%	
	M3	QT	50%	
	M4	QT	95%	
Timeliness	M1	QL	Não se aplica	X
	M2	QL	Não se aplica	
Índice global de qualidade	66,62%			

Fonte: Autoria própria (2018).

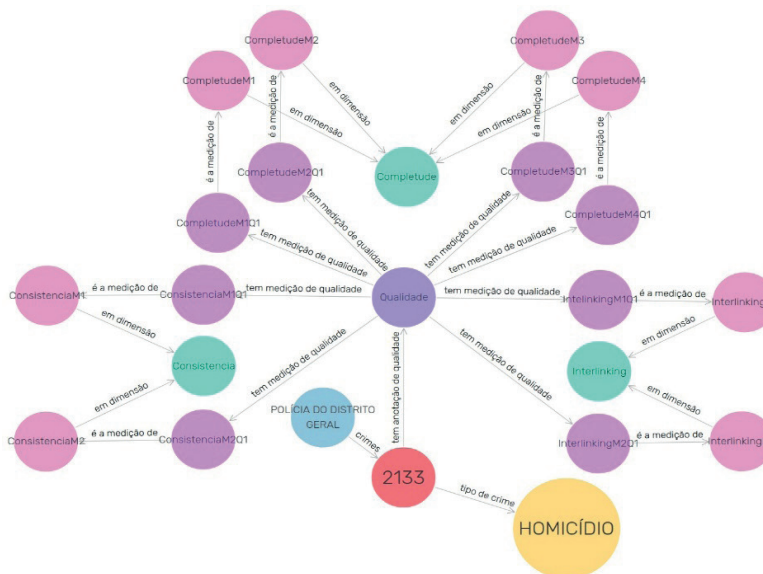
A Figura 1 representa a instância de um dos recursos presentes no conjunto de dados analisado (crime 2133), vinculada à instância de qualidade por meio da propriedade "tem anotação de qualidade". Esta instância foi criada com o uso do vocabulário *Data Quality Vocabulary* (DQV), estabelecendo as dimensões e respectivas métricas de acordo com a metodologia apresentada.

O DQV, em português, Vocabulário de Qualidade de Dados, é um modelo proposto pela W3C como extensão ao *Data Catalog Vocabulary* (DCAT), em português, Vocabulário do Catálogo de Dados, com foco em possibilitar a representação da qualidade dos dados, principalmente conectados, disponibilizados em repositórios (W3C, 2016).

Este vocabulário conta com um conjunto de classes e propriedades, seguindo os padrões de LOD e Web Semântica, que possibilitam a elaboração de esquemas para representação de diversas dimensões, métricas e valores relacionados aos dados, sendo possível desenvolvê-los conforme o domínio ou características do conjunto de dados. Possui a capacidade de representar a qualidade tanto do dado propriamente dito, quanto a relacionada a metadados (W3C, 2016).

Isso torna muito atraente a aplicação desse vocabulário em conjuntos de dados abertos conectados. Além de conseguir representar a qualidade dos dados contidos no repositório, é possível avaliar questões sobre os padrões e metadados relacionados ao repositório, o qual foi abordado neste capítulo.

Figura 1 – Representação das métricas de qualidade aplicadas a um recurso presente no conjunto de dados avaliado.



Fonte: Autoria própria (2018).

4 Considerações finais

Este trabalho propôs-se a apresentar uma avaliação de qualidade com base em um modelo de qualidade de dados para o domínio de dados conectados, dividido em três pilares:

a) literatura, por meio da qual se obtiveram informações sobre problemas e dimensões de qualidade para dados conectados;
b) W3C, que estabelece os padrões para o funcionamento tanto da Web Semântica, como dos dados conectados;
c) projeto *Linked Open Data* (LOD), o qual estabelece princípios de qualidade, reúne conjuntos de dados, organiza-os em diferentes categorias e promove a visibilidade dos que atendem a tais princípios.

Por meio do modelo foi utilizada a seguinte metodologia de avaliação:

a) levantamento de requisitos de qualidade para os dados conectados;
b) definição das dimensões e métricas;
c) avaliação de qualidade.

Foram definidas as seguintes dimensões de qualidade: *interlinking*, consistência, completude, licenciamento, avaliação temporal, precisão sintática e semântica. Bem como métricas de avaliação para os problemas específicos de cada dimensão, utilizando 14 fórmulas para realizar uma avaliação quantitativa do conjunto de dados.

Ao realizar o processo de avaliação, os seguintes fatos puderam ser constatados:

a) é muito importante, no processo de avaliação de qualidade, que o avaliador tenha um conhecimento considerável, ou procure compreender os vocabulários utilizados para descrição dos dados, visto que cinco das sete dimensões propostas fazem uso de tais informações para realizar a avaliação;
b) o conjunto avaliado se mostrou falho em relação aos requisitos de qualidade para dados conectados e web semântica em dois requisitos de *interlinking*. O primeiro, não utilizar *namespaces* que fazem uso de conjuntos ou

identificadores numéricos para os recursos, quando utilizados, as URIs não podem ser referenciáveis, visto que apenas o domínio que as criou conseguiria identificá-las. Uma URI é referenciável quando é possível recuperar o recurso identificado por ela. O segundo, utilizar chaves naturais nas URIs: utilizar identificadores significativos para o domínio do conjunto de dados, como a combinação do nome e sobrenome do recurso (no caso de uma pequena organização). Tais requisitos podem ser considerados de extrema importância, visto que são amplamente disseminados na comunidade de boas práticas para web semântica e dados conectados;

c) quanto à disponibilização da informação, as métricas 1, 2 e 3 de dimensão de completude mostram que apenas 30,43% das classes e propriedades disponíveis no vocabulário criado especificamente para descrição de recursos criminais foram utilizados, levando à conclusão de que mais de 60% de informações que poderiam ser disponibilizadas não o foram.

A qualidade tem sido muito estudada na área da Ciência da Computação, entretanto, entende-se que hoje em dia a publicação de dados é algo assumido pela Ciência da Informação. Simultaneamente, em ambas as áreas, existe a necessidade da devida representação dos dados de qualidade publicados para que sejam legíveis e compreensíveis, tanto por máquinas quanto por humanos, possibilitando interpretações mais amplas. Tais interpretações podem servir de insumo para que novas informações possam ser formuladas, permitindo assim que novas métricas de qualidade sejam utilizadas e posteriormente representadas, tornando o processo de avaliação e representação de qualidade iterativo e mais ágil, pela facilidade de visualização e interação com dados de qualidade.

Sendo assim, é imprescindível que a publicação de dados seja realizada com boa qualidade. A metodologia apresentada contribui nesse contexto, para avaliar e revelar a qualidade de dados, favorecendo processos informacionais relacionados à gestão da informação e consequentemente a tomada de decisão em ambientes diversos. Acredita-se que a metodologia possa ser utilizada para detectar problemas de qualidade nos processos de armazenamento, recuperação e

manipulação dos dados e proporcionar que humanos e sistemas possam usufruir de melhores subsídios informacionais.

Referências

BATINI, C.; SCANNAPIECO, M. **Data and information quality**: dimensions, principles and techniques. [S. l.]: Springer, 2016.

BERNERS-LEE, T. **Cool URIs don't change**. 2008. Disponível em: https://www.w3.org/Provider/Style/URI. Acesso em: 8 jun. 2018.

BERNERS-LEE, T. **Linked data**. 2006. Disponível em: http://www.w3.org/DesignIssues/LinkedData.html. Acesso em: 8 jun. 2018.

BIZER, C.; CYGANIAK, R. Quality-driven information filtering using the wiqa policy framework. **Journal of Web Semantics**, v. 7, n. 1, p. 1-10, 2009. Disponível em: https://www.sciencedirect.com/science/article/abs/pii/S157082680800019X. Acesso em: 8 de jun. 2018.

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

BORKO, H. Information science: what is it? **Journal of the Association for Information Science and Technology**, v. 19, n. 1, p. 3-5, 1968.

BOUZEGHOUB, M. A framework for analysis of data freshness. *In*: INTERNATIONAL WORKSHOP ON INFORMATION QUALITY IN INFORMATION SYSTEMS, 2004, Paris. **Proceedings** [...]. New York: ACM, 2004.

HEATH, T.; BIZER, C. Linked data: Evolving the web into a global data space. **Synthesis lectures on the semantic web: theory and technology**, v. 1, n. 1, p. 1-136, 2011.

KONTOKOSTAS, D. et al. Triplecheckmate: a tool for crowdsourcing the quality assessment of linked data. *In*: INTERNATIONAL CONFERENCE ON KNOWLEDGE ENGINEERING AND THE SEMANTIC WEB, 4., 2013, St. Petersburg. **Proceedings** [...]. St. Petersburg: Springer, 2013. p. 265-272. Disponível em: https://link.springer.com/chapter/10.1007/978-3-642-41360-5_22. Acesso em: 8 jun. 2018.

MENDES, P. N.; MÜHLEISEN, H.; BIZER, C. Sieve: linked data quality assessment and fusion. *In*: JOINT EDBT/ICDT WORKSHOPS, 2012, Berlin. **Proceedings** [...]. Berlin: ACM, 2012. p. 116-123. Disponível em: https://dl.acm.org/citation.cfm?id=2320803. Acesso em: 8 jun. 2018.

RULA, A.; ZAVERI, A. Methodology for assessment of linked data quality. *In*: WORKSHOP ON LINKED DATA QUALITY, 1., 2014, Leipzig. **Proceedings** [...]. Leipzig: [s.n.], 2014. Disponível em: http://ceur-ws.org/Vol-1215/paper-04.pdf. Acesso em: 8 jun. 2018.

SAUERMANN, L.; CYGANIAK, R.; VÖLKEL, M. Cool uris for the semantic web. **Sonstige Einrichtungen**, v. 7, n. 1, 2007.

SOUSA, J. O. **Metodologia de avaliação de qualidade de dados no contexto do linked data**. 2017. Dissertação (Mestrado em Ciência da Informação) – Universidade Estadual Paulista, Marília, 2017.

W3C. **Data on the web best practices**: data quality vocabulary. 2016. Disponível em: https://www.w3.org/TR/vocab-dqv/. Acesso em: 9 jun. 2018.

ZAVERI, A. *et al*. Quality assessment for linked data: a survey. **Semantic Web**, v. 7, n. 1, p. 63-93, 2015.

ZAVERI, A. et al. User-driven quality evaluation of DBpedia. In: INTERNATIONAL CONFERENCE ON SEMANTIC SYSTEMS, 9., 2013, Graz. **Proceedings** [...]. Graz: ACM, 2013.

Ansiedade da informação revisitada: reflexões teóricas com base na psicologia e na ciência da informação

Henry Poncio Cruz de Oliveira
Josevânia da Silva

1 Debate inicial

A sociedade contemporânea, marcada pela produção, disseminação, acesso e uso de vastos volumes de informação de natureza digital, foi denominada pelo sociólogo polonês Bauman (2001) de modernidade líquida. Recorrendo ao que a física postula sobre a estrutura e os estados físicos da matéria, Bauman (2001) utiliza, metaforicamente, a ideia da liquidez para, de forma crítica, redimensionar a noção de emancipação, de tempo e espaço, de trabalho, de comunidade, de consumo, de amor. O autor afirma serem inegáveis, no contexto da modernidade líquida, as profundas mudanças ocorridas no convívio social e nos comportamentos humanos e destaca o papel das Tecnologias de Informação e Comunicação nestas mudanças.

Trata-se de um paradigma informacional e tecnológico capitalizado em todos os setores da sociedade. Este paradigma promove implicações nas subjetividades, nas formas de construção de relações com o outro e na capacidade de atribuir significado aos eventos do cotidiano.

No contexto da modernidade líquida, a Organização Mundial de Saúde tem advertido sobre o crescimento dos casos de transtornos mentais. Os dados do relatório *"Depression and Other Common Mental Disorders: Global Health Estimates"* sinalizam que a depressão é classificada como o maior contribuinte individual para a incapacidade (7,5%), já os transtornos de ansiedade estão em 6º lugar (3,4%) no ranking global.

As consequências desses distúrbios, em termos de perda de saúde, são significativas. No Brasil, 8,3 % da população sofre com transtornos de ansiedade

e 10,3 % com transtornos depressivos (WORLD HEALTH ORGANIZATION, 2017). Os transtornos de ansiedade podem ser desencadeados por aspectos multifatoriais (DAVIDOFF, 2001), dentre os quais podemos destacar os fatores informacional e tecnológico como potenciais para o comportamento ansioso.

Neste capítulo, apresentamos o interesse especial de revisitar a noção de ansiedade relacionada à informação para gerar uma agenda de pesquisa sobre o fenômeno da ansiedade da informação. Para tanto, o aporte metodológico utilizado neste trabalho científico é a ensaística. Rodríguez (2012) defende o ensaio como uma alternativa metodológica nas Ciências Humanas e Sociais, que possibilita romper com o objetivismo puro.

A ensaística compreende uma estruturação narrativa marcada por uma "composição textual argumentativa que permite enunciar elementos concretos e abstratos com suficiente conflito e transformação de ideias, podendo completá-las ou delas duvidar" (Rodríguez, 2012, p. 92). Do ponto de vista analítico, é um método qualitativo que permite a produção de textos científicos considerando a experiência autoral.

Incialmente, são apresentados aspectos conceituais sobre o construto ansiedade, seguido da explanação sobre a ansiedade de informação. Por último, é apresentada uma agenda de pesquisa para o estudo sobre a ansiedade de informação.

2 Ansiedade: aspectos conceituais

A vivência de momentos perpassados por estados de ansiedade é comum à condição humana durante seu desenvolvimento. Em geral, situações inesperadas ou que, aparentemente, fogem ao controle dos sujeitos podem gerar níveis de ansiedade. Dada a frequência com que as pessoas vivenciam a ansiedade, o uso do termo é empregado, muitas vezes, como sendo sinônimo de medo ou stress. Não obstante, trata-se de construtos diferentes.

O medo é caracterizado como umas das emoções básicas, de natureza primitiva, que se manifesta em função de alguma ameaça presente ou iminente, caracterizada por significativa excitação e tendência à ação (BARLOW, 2002).

Em termos conceituais, é definido como "um estado neurofisiológico automático primitivo de alarme envolvendo a avaliação cognitiva de ameaça ou perigo iminente à segurança e integridade de um indivíduo" (CLARK; BECK, 2012, p. 17).

Já o stress, ou mais especificamente o estressor, segundo Towsend (2011, p. 584), "é uma pressão externa que é exercida sobre o indivíduo" que pode gerar impactos transitórios ou permanentes. Embora o termo seja, frequentemente, associado a aspectos negativos, é importante dizer que o stress pode gerar respostas positivas e adaptativas (eustresse), bem como provocar sobrecarga e prejuízo para as pessoas, sendo este último denominado distresse (NODARI *et al*., 2014).

A ansiedade, por sua vez, é caracterizada como um sistema de resposta cognitiva, afetiva, fisiológica e comportamental complexo, manifestado em decorrência da antecipação de eventos considerados aversivos, sobretudo pelo seu caráter imprevisível e incontrolável para os sujeitos, os quais são compreendidos como ameaça aos interesses da pessoa que sente (CLARK; BECK, 2012). Alguns autores a definem como "uma apreensão difusa que é vaga na sua natureza e está associada com sentimentos de incerteza e impotência" (TOWSEND, 2011, p. 17), bem como "uma emoção orientada ao futuro, caracterizada por percepções de incontrolabilidade e imprevisibilidade" (CLARK; BECK, 2012, p. 104).

O medo pode estar presente em transtornos de ansiedade, como o Transtorno de Pânico. Contudo, segundo Clark e Beck (2012), a ansiedade refere-se a um estado mais permanente de ameaça e pode abarcar outras emoções como incerteza, sentimentos de desamparo e de incapacidade para conseguir atingir os resultados almejados (BARLOW, 2002). Nesta direção, a vivência da ansiedade é idiossincrática, tendo em vista que o significado sobre o evento ameaçador varia de pessoas para pessoa, conforme os modos de enfrentamento de cada indivíduo e vivências anteriores.

Na contemporaneidade, os indicadores epidemiológicos de ansiedade têm aumentado nas últimas décadas. A literatura aponta que o nível de ansiedade vivenciada por crianças, na atualidade, equipara-se ao mesmo nível de ansiedade de pacientes com diagnósticos psiquiátricos da década de 1950 (LEAHY, 2012). Em parte, o aumento nos níveis de ansiedade das sociedades está relacionado com as desigualdades sociais, com a escassez de políticas públicas que promovam

e garantam o bem-estar social dos indivíduos e populações, incluídos aí aspectos relativos ao emprego, à alimentação, à segurança, às condições de moradia, dentre outros aspectos.

Além das condições objetivas de vida, a literatura aponta como geradores de ansiedade a fragilidade e a fluidez das relações sociais (LEAHY, 2012) e a busca incessante por informações ou por estar-se sempre "atualizado", o que Richard Saul Wurman denominou de "ansiedade de informação" (WURMAN, 1991). Para este autor, a ansiedade de informação decorreria da percepção dos atores sociais sobre a distância entre o que eles acreditam saber e o que eles poderiam saber. No contexto da Teoria Cibernética, fala-se de um sentimento de emergência por informação que, muitas vezes, não corresponde à capacidade dos indivíduos de elaborarem ou assimilarem-na, o que pode gerar ansiedade e a sensação de não estarem incluídos (SIQUEIRA, 2017).

Mediante o exposto, se faz necessário discorrer sobre o que vem a ser a chamada "ansiedade de informação". Cabe, ainda, problematizar em que medida a procura por informação, às vezes de caráter meramente quantitativo, repercute no modo como indivíduos e coletividades interagem e elaboram sentidos sobre os fenômenos sociais.

3 Ansiedade de informação

Temos vivido os desdobramentos de uma revolução tecnológica que, a cada dia, consolida-se em um paradigma sociotécnico de permanente construção, desconstrução e reconstrução dos elementos constitutivos da vida em sociedade. Dentre os mais significativos elementos desta revolução tecnológica está a internet, que, segundo Wurman (2005), transformou nosso estilo de vida, contribuiu para a formação as novas maneiras de fazer negócios e produziu crises por meio das tecnologias emergentes.

A internet tem propiciado a conexão de pessoas e organizações em uma estrutura de rede onde são produzidas e disseminadas massas de dados e informações de natureza digital. As massas de dados e informações supracitadas têm sido armazenadas em servidores e *data centers* espalhados em todo o globo.

Além disso, os dados e informações circulam através de ambientes de informação estruturados em teias hipertextuais, o que contribui para a construção de conhecimento. Nessa direção, Oliveira (2014, p. 87) assevera que:

> Nesse entrelaçado de modificações sociotécnicas, o crescente quantitativo de informações criou necessidades de reflexão e revisão das formas de apresentação da informação considerando a espacialidade que ela ocupa em ambientes digitais e não digitais.

Na contemporaneidade, além do melhoramento nas tecnologias de armazenamento e divulgação, ocorreu a ampliação do processo de democratização do acesso à informação. Todavia, para os sujeitos, tais melhorias não correspondem a uma melhor qualidade na elaboração, reflexão e construção de conhecimento a partir da infinidade de informações disponíveis. Sobretudo porque os volumes de informação e de não-informação gerados cotidianamente têm produzido impacto nos sujeitos, gerando sintomas de ansiedade (OLIVEIRA, 2014).

Wurman (1991, p. 49) destaca que todos apresentam algum grau de ansiedade de informação. Diz ainda que ela pode "nos atingir em qualquer nível e pode resultar tanto de excesso quanto de falta de informação". Ainda, segundo este autor, existem vários contextos que podem provocar a ansiedade de informação, como: "não compreender a informação; sentir-se assoberbado por seu volume; não saber se uma certa informação existe; não saber onde encontrar a informação; saber exatamente onde encontrá-la e não conseguir acesso". (WURMAN, 1991).

A ansiedade de informação pode gerar sofrimento psíquico, bem como desencadear mudanças comportamentais. Wurman (1991) destaca uma variedade de comportamentos que, dependo do nível de ocorrência, caracterizam-se como sintomas indicativos de ansiedade de informação. Os comportamentos supracitados estão agrupados no Quadro 1.

Quadro 1 – Comportamentos indicativos de Ansiedade de Informação

Falar compulsivamente que não consegue se manter atualizado com o que ocorre ao seu redor.
Sentir-se culpado pelo volume de informações à espera de leitura.
Balançar a cabeça compenetradamente quando alguém menciona uma informação ou notícia que nunca tinha ouvido falar.
Descobrir que é incapaz de explicar algo que pensava ter entendido.
Xingar-se por não ser capaz de seguir as informações contidas em manuais.
Recusar-se a comprar um equipamento eletrônico por medo de não conseguir operá-lo.
Sentir-se deprimido por não saber para que servem os botões e ícones de um equipamento eletrônico.
Comprar equipamentos eletrônicos de alta tecnologia achando que pode aprender a tecnologia por osmose.
Qualificar um texto como "genial" mesmo sem ter compreendido seu resumo ou resenha.
Dedicar tempo e atenção a notícias que não têm qualquer impacto cultural, econômico ou científico na sua vida.
Ao preencher um formulário, sentir-se obrigado a ocupar todos os espaços em branco.
Reagir emotivamente à informação que você de fato não compreende.
Achar que a pessoa ao seu lado está compreendendo tudo e você não.
Sentir-se incomodado por ter que dizer "não sei".
Chamar de informação alguma coisa que você não compreende.

Fonte: Adaptado de Wurman (1991, 2005)

A informação se transformou na força motriz de nossas vidas e melhorou a qualidade de vida das sociedades em vários aspectos, como, por exemplo, a melhoria das tecnologias voltadas aos cuidados em saúde. Contrariamente, dependendo do modo como os indivíduos lidam com o volume de informações de que dispõem, o excesso de informação pode se caracterizar como ameaça, sobretudo quando demanda esforços cognitivos significativos para compreensão, estando tal esforço associado a um sentimento de não conseguir abarcar tudo que deseja, o que contribui para a ansiedade de informação.

> A ansiedade de informação é causada pela distância cada vez maior entre o que compreendemos e o que achamos que deveríamos compreender. É o buraco negro existente entre os dados e o conhecimento, que aparece quando a informação não diz o que queremos saber. (WURMAN, 2005, p. 14).

O autor supracitado assevera que as pessoas permanecem sentindo ansiedade em relação a como assimilar um corpo de informações que se modifica e se expande a cada nanossegundo. Contraditoriamente, neste volume crescente de informações há um significativo volume de não-informação, desinformação e confusão. Shedroff (2005) presume que, embora nosso cérebro tenha uma capacidade de armazenar e processar mais dados do que os que nós utilizamos, é certo que a tentativa de ficar cientes de tudo que nos rodeia produz uma ansiedade que nos invade e nos afeta mental, emocional e até fisicamente.

A ansiedade de informação pode se mostrar de várias formas. Em primeiro, a ansiedade pode decorrer da frustração gerada pela incapacidade de conseguirmos ficar a par de tudo que ocorre em nosso entorno (SHEDROFF, 2005). Em segundo, a ansiedade também decorre da velocidade com que as informações são apresentadas. Os dados e informações não são passivos, são dinâmicos, infiltram-se constantemente em nossos ambientes e invadem nossa atenção. Nesse contexto, ocorre um "sentimento de culpa por não estarmos "mais informados" e não sermos capazes de acompanhar o volume de dados considerados informação." (SHEDROFF, 2005, p. 16).

Uma terceira forma de ansiedade de informação, mais sutil, materializa-se na "[...] frustração diante da qualidade do que nos é oferecido – sobretudo do que se apresenta como notícia" (SHEDROFF, 2005, p. 16). Assim, a tríade composta pela "necessidade de estar sempre informado", pela "velocidade de disseminação das informações" e pela "qualidade das informações noticiadas" dinamiza-se no conflito.

Shedroff (2005) apresenta uma quarta via de manifestação da ansiedade de informação, que se associa ao que o autor chama de "arrogância de saber antes dos outros". Trata-se de um comportamento ansioso que remete o sujeito a uma necessidade de informar-se antecipadamente para que, ao ser questionado por

um terceiro sobre um tema específico, não seja surpreendido por desconhecer o que circula em relação a este tema específico.

É possível dizer que a ansiedade construída frente à necessidade de informar-se na grande massa de dados e informações, volatilizadas pela cultura tecnológica da liquidez, incide direta e negativamente no cuidado pela qualidade, fidedignidade, segurança e autenticidade da informação. Ansiosos para estar informados, consolidamos comportamentos que priorizam números e quantidades ao invés de qualidade, cliques e curtidas ao invés de conteúdo relevante.

Uma questão latente, para Wurman (1991), é a aceitação do caos informacional causado pelo volume crescente de informação disponível. Para este autor, compreender e aceitar a complexidade do fenômeno informacional contemporâneo pode ser uma estratégia para a redução da ansiedade de informação.

No cerne do debate sobre ansiedade de informação está a necessidade de compreender a não-informação. No contexto da era da informação, marcada pela liquidez, existe uma explosão de não-informação, ou seja, uma explosão de dados que não necessariamente conduzem as pessoas à construção de conhecimento.

O enfrentamento desta questão, segundo Wurman (1991), começa pela distinção que deve ser feita entre dados e informação. Para Wurman (1991), a informação se efetiva num processo cognitivo que conduz o sujeito à compreensão de uma realidade. Já os dados têm o potencial de estruturarem informação, mas à priori não conduzem o sujeito a um processo cognitivo que culmina com a construção ou reconstrução de um conhecimento.

A grande aposta de Wurman (1991, 2005) para a redução da ansiedade da informação é a Arquitetura da Informação. Este autor defende que a provisão do acesso adequado possibilita que o sujeito encontre a informação de que precisa. Estratégias para organização e representação da informação impactam diretamente na recuperação da informação, possibilitando o uso adequado e a redução da ansiedade de informação, visto que "a informação pode ser infinita, mas não as formas de estruturá-la". Assim, descobrir princípios de organização "é tão essencial para você trabalhar com os conjuntos de informação quanto para desenvolver seus programas pessoais de informação" e "compreender a estrutura e a organização da informação permite extrair dela valor e significado".

4 Ansiedade de informação: ensaiando uma agenda de pesquisa

Na exposição que fizemos nas seções 2 e 3 sobre ansiedade e ansiedade da informação, respectivamente, procuramos estruturar uma trama conceitual que sustentasse a atividade ensaística de gerar uma agenda de pesquisa em torno da ansiedade da informação.

A construção teórica e conceitual sobre ansiedade da informação foi feita, propositalmente, priorizando as obras Ansiedade da Informação (1991) e Ansiedade da Informação 2 (2001), do arquiteto Richard Soul Wurman. A escolha se justifica por este autor ser pioneiro na percepção empírica e contribuição conceitual do fenômeno da ansiedade da informação.

É necessário refletir sobre as direções que Wurman (1991, 2005) aponta para tratar o fenômeno da ansiedade da informação. Conforme apresentamos, Wurman (1991, 2005) aposta em soluções prioritariamente exógenas, ou seja, a resolução ou pelo menos a redução da ansiedade de informação, que é um fenômeno fundamentalmente humano e subjetivo, pode ser alcançada de forma pragmática e objetiva à medida que se investe na organização da informação, racionalizando sua estrutura por meio de dimensões arquiteturais.

É irrefutável que a Arquitetura da Informação é capaz de reduzir incertezas informacionais, reduzir sobrecarga cognitiva e desorientação espaço-temporal e minimizar o sentimento de ansiedade (OLIVEIRA, 2014). Porém, aproximadamente duas décadas depois das publicações de Wurman (1991, 2005), se faz necessário o incremento em estudos e pesquisas que considerem minimizar a ansiedade de informação, tendo como estratégia, por exemplo, pesquisas no contexto da arquitetura da informação.

Os elementos característicos da modernidade líquida e os indicadores epidemiológicos apontam para esta necessidade, conforme demonstrado através dos dados globais e brasileiros sobre ansiedade, apresentados no relatório *"Depression and Other Common Mental Disorders: Global Health Estimates"*. Isto exige atentarmos para o paradigma sociotécnico marcado pela produção e uso intenso de informação e tecnologias.

A retomada dos estudos e pesquisas sobre ansiedade da informação é ainda mais necessária quando se considera os volumes e estoques de dados e informações atuais, os quais são significativamente maiores que os da década de 1990 e começo dos anos 2000. Ademais, no período mencionado, as tecnologias móveis e as redes sociais não eram uma realidade tão marcante, o que nos permite inferir que a ansiedade da informação pode conter novos elementos constitutivos e substanciais, que merecem ser investigados.

Diante do exposto, apresentamos uma agenda de pesquisa com foco no fenômeno da ansiedade da informação, estruturada por meio da construção de uma pergunta de pesquisa e do delineamento de um objetivo de pesquisa, conforme exposto no quadro 02.

Quadro 2 – Agenda de Pesquisa sobre Ansiedade de Informação

Questão de Pesquisa	**Objetivo**
No contexto da modernidade líquida, como podemos (re)conceituar ansiedade de informação?	Elaborar um quadro conceitual para ansiedade de informação que considere a cultura da fluidez e liquidez.
Existem instrumentos de pesquisa para mensurar a ansiedade de informação nos dias atuais?	Realizar pesquisa metodológica para construção de instrumentos de mensuração da ansiedade de informação.
Os elementos comportamentais indicativos de Ansiedade de Informação encontrados em Wurman (1991, 2005) podem ser insumo para um instrumento de avaliação de ansiedade de informação?	Construir um instrumento de avaliação de ansiedade de informação com base nos elementos comportamentais indicativos de Ansiedade de Informação de Wurman (1991, 2005).
Qual a relação entre Ansiedade de Informação e redes sociais?	Investigar relações entre Ansiedade de Informação e redes sociais.
Qual a relação entre Ansiedade de Informação e dispositivos móveis?	Investigar relações entre Ansiedade de Informação e dispositivos móveis.
Qual a relação entre Ansiedade de Informação e experiências *cross-channel*?	Investigar relações entre Ansiedade de Informação e experiências *cross-channel*.
Quais as diferenças entre Ansiedade de Informação e stress informacional?	Diferenciar Ansiedade de Informação e stress informacional?
Como enquadrar a Ansiedade de Informação no contexto dos Transtornos de Ansiedade?	Categorizar a Ansiedade de Informação no contexto dos Transtornos de Ansiedade.

Fonte: Autoria própria (2018).

5 Considerações finais

Acreditamos que o exposto como agenda de pesquisa para o fenômeno da Ansiedade de Informação necessita de aparatos teóricos e metodológicos advindos das áreas de Ciência da Informação e da Psicologia.

Por conseguinte, compreendemos que o fenômeno da Ansiedade de Informação é complexo e multifatorial, necessitando de pesquisas do tipo interdisciplinar e de caráter qualitativo predominante.

O presente texto, enquadrado como um ensaio teórico, tratou de resgatar a Ansiedade de Informação enquanto fenômeno de investigação de interesse dos campos científicos informacional e psicológico. Nossa pretensão em gerar demandas de pesquisa a partir da problematização da ansiedade da informação enquanto categoria científica tem o potencial de contribuir na compreensão das nuances que este fenômeno adquire nos dias atuais, dias cheios de informação, repletos de telas e toques e de linguagem imagética que circula em todas as direções e sentidos da rede.

Referências

BAUMAN, Z. **Modernidade líquida**. Rio de Janeiro: Jorge Zahar Editor, 2001.

BARLOW, D.H. **Anxiety and its disorders**: the nature and treatment of anxiety and panic. New York: The Guilford Press, 2002.

CLARK, D. A.; BECK, A. T. **Terapia cognitiva para os transtornos de ansiedade**: ciência e prática. Porto Alegre: Artmed, 2012

DAVIDOFF, L. L. **Introdução à psicologia**. São Paulo: Perason Makron Books, 2001.

LEAHY, R. L. **Livre de ansiedade**. São Paulo: Artmed Editora, 2012.

NODARI, N. L. *et al*. Estresse, conceitos, manifestações e avaliação em saúde: revisão de literatura. **Saúde e Desenvolvimento Humano**, v. 2, n. 1, p. 61-74, 2014.

OLIVEIRA, H. P. C. de. **Arquitetura da informação pervasiva**: contribuições conceituais. 2014. 202 F. Tese (Doutorado em Ciência da Informação) – Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2014.

RODRÍGUEZ, V. G. **O ensaio como tese**: estética e narrativa na composição do texto científico. São Paulo: WMF Martins Fontes, 2012.

SHEDROFF, N. Formas de ansiedade de informação. *In*: WURMAN, R. S. **Ansiedade de informação 2**: um guia para quem comunica e dá instruções. São Paulo: Cultura, 2005.

SIQUEIRA, T. G. S. Reflexiones sobre la necesidad o importancia de la información en la contemporaneidad. **Revista Prefácio**, v. 1, n. 1, p. 31-38, 2017.

TOWSEND, M. **Enfermagem em Saúde Mental e Psiquiátrica**: conceitos de cuidado na prática baseada na evidência. Loures: Lusociência, 2011.

WORLD HEALTH ORGANIZATION. **Depression and other common mental disorders**: global health estimates. Geneva: WHO Document Production Services, 2017.

WURMAN, R. S. **Ansiedade de informação**. São Paulo: Cultura Editores Associados, 1991.

WURMAN, R. S. **Ansiedade de Informação 2**. São Paulo: Editora de Cultura, 2005.

El lenguaje de la documentación en salud: algunos obstáculos evitables[9]

Rosa Estopà Bagot

1 Presentación

Nuestra vida está rodeada de palabras. Desde que nacemos hasta que nos morimos, las palabras, a veces conscientemente y otras no tanto, nos acompañan siempre. El *lenguaje* no es una posibilidad ni una elección, sino una **necesidad** inherente a la especie humana. Nuestra cotidianidad está totalmente unida al hecho lingüístico, es como han defendido algunos lingüistas (TUSÓN, 2002) la **marca distintiva de la humanidad**. También los entornos profesionales y especializados necesitan lenguaje para representar y comunicar su saber. La salud no es una excepción.

Los avances científicos y tecnológicos de los últimos tiempos han acelerado notablemente la evolución de la humanidad en todos los ámbitos, incluido el relativo a las ciencias de la salud. Sin embargo, la mayoría de especialistas (KURTZ; 2002; GALARZA *et al*., 2003; MOORE, 2010; SHUKLA; YADAV; KASTURY, 2010) coinciden en afirmar que en la comunicación médico-paciente todavía es muy necesario continuar avanzando.

Entre los profesionales de las ciencias de la salud cada vez se extiende más la constatación de que la capacidad de comunicarse de manera efectiva con los pacientes es crucial para la eficacia del tratamiento que se lleve a cabo. Y ya empieza a ser habitual que los programas de formación de dichos profesionales proporcionen técnicas de comunicación eficaces, que se conciben como herramientas terapéuticas adicionales. Con ello se pretende mejorar la eficacia del tratamiento terapéutico

9 Este trabajo se ha desarrollado en el marco del proyecto JUNTS: superar barreres socioeducatives i afavorir l'alfabetització sobre les interferències i dificultats de comprensió de la informació i documentació dirigida a famílies d'infants afectats per malalties rares. RecerCaixa 2015. Avancem amb la ciència. ACUP y Obra Social "La Caixa". IP: Rosa Estopà Bagot (Universitat Pompeu Fabra) y Manuel Armayones (Universitat Oberta de Catalunya). Disponible a: https://www.upf.edu/web/medicina_comunicacio.

y, por ende, el estado de salud del paciente, pero también incidir positivamente en el bienestar del profesional de la salud, con el fin de mejorar su satisfacción y su sentido de la eficacia profesional.

Desde el punto de vista del paciente, en las últimas décadas se está produciendo un cambio importante gracias a potenciar su empoderamiento (SØRENSEN *et al*., 20015), en virtud del cual el paciente está dejando de ser pasivo para convertirse en una persona activa que toma parte de las decisiones sobre su salud. De hecho, las políticas sanitarias actuales abogan por este empoderamiento del paciente a través de programas de *alfabetización en salud*. Este complejo concepto – traducción del término inglés *health literacy* – comenzó a introducirse en la década de los 70 y puede definirse, siguiendo el glosario de términos de salud pública de la Organización Mundial de la Salud (OMS), como "las habilidades sociales y cognitivas que determinan el nivel de motivación y la capacidad de una persona para acceder, entender y utilizar la información de forma que le permita promover y mantener una buena salud" (FALCÓN; LUNA, 2012, p. 92).

Uno de los factores fundamentales para aumentar la alfabetización en salud es asegurar una buena comunicación entre el personal médico y el paciente. Sin embargo, es sabido que a menudo dicha comunicación no es satisfactoria, y que los pacientes no siempre comprenden adecuadamente los actos comunicativos, orales y escritos, producidos por los profesionales de la salud, ya que existen brechas de tipo lingüístico y cognitivo que dificultan la comprensión de los enunciados (DELÀS, 2005; BORRELL, 2004). Estos problemas de comprensión acarrean consecuencias psicosociales y sanitarias negativas, puesto que es indispensable que exista una buena comunicación con el profesional de la salud para conseguir una alfabetización en salud inclusiva, en la que el paciente sea un agente activo y colaborador.

2 El léxico de la medicina

La medicina y todas las ciencias de la salud necesitan léxico con valor especializado para hablar de la salud y de la enfermedad, de los tratamientos y de las curas, de los nuevos avances y de la anatomía, etc. El léxico de la medicina es el

resultado de dos mil quinientos años de pensamiento científico y se fundamenta en lenguas clásicas que en la antigüedad han tenido mucha influencia en el desarrollo de la medicina: en el griego, el árabe y el latín (LÓPEZ PIÑERO; TERRADA FERRANDIS, 1990; BERNABEU *et al*., 1995; GUTIÉRREZ, 1998). Pero sobre todo la terminología médica se fundamenta en un conjunto reducido de formantes grecolatinos. Así, unos 1000 formantes clásicos (como, por ejemplo, *-itis*, que significa 'inflamación'; *cardio-*, 'corazón'; *taqui-*, 'rápido'; *-soma*, 'cuerpo', *polio-*, 'gris'; *-osis*, 'estado patológico'; *hepato-*, 'hígado'; *-oma*, 'tumor') permiten generar el 80 % del léxico de las ciencias de la salud (ANDERSON, 2016): *hepatitis*, *cardiopatía*, *taquicardia*, *cromosoma*, *polioencefalitis*, *varicosis*, *anquilostoma*, *hematoma*, *hepatoma*, *diagnosis*, etc.

Una gran cantidad de los términos médicos (en abstracto más del 80%, en contexto la cifra se reduce considerablemente) son términos poliléxicos, es decir formados por más de una palabra. Los sintagmas formados por un nombre y un adjetivo o un sintagma preposicional son los más usuales. Estas dos estructuras son las más frecuentes porque permiten especificar el núcleo del sintagma, permiten especificar y delimitar. Las lenguas románicas utilizan esta estrategia morfosintáctica (los adjetivos) para especificar un nombre y tipologizarlo: *hepatitis es el genérico, pero cuando queremos ser más específicos utilizamos términos como hepatitis A, hepatitis C, hepatitis alcohólica, hepatitis colestática, hepatitis crónica activa, hepatitis crónica autoimmunitaria, hepatitis infecciosa, hepatitis lúpica, hepatitis sérica, hepatitis tóxica, hepatitis transfusional, hepatitis vírica*, etc.

Más allá de los términos creados con formantes clásicos (que son la mayoría), actualmente el léxico de la salud también está influenciado por lenguas como el alemán y el francés, pero sobre todo por el inglés. Así, palabras como *anding, buffer, bypass, flap, flapping tremor, flutter, follow-up, lamping, core, distress, doping, feed-flush, flutter, handicap, immunoblotting, killer, mapping, odds ratio, patch test, pool, output, shunt, rash, relax, scanner, score o screening* (NAVARRO, 2001; GARCÍA PALACIOS, 2004) son solo una pequeña muestra de la gran cantidad de términos nuevos que llegan al español directamente de la lengua inglesa. Además, estos términos suelen ser unidades relacionadas con la tecnología y las nuevas técnicas y aparatos para realizar pruebas y tratamientos más precisos.

Otra de las características del léxico de la medicina es el uso de los epónimos. Así los sintagmas en los que el núcleo del sintagma preposicional especificador es un nombre propio son muy abundantes. Cuando se trata de un antropónimo (normalmente el apellido de una persona), suelen indicar el médico que describió por primera vez un fenómeno o uno de los primero pacientes reconocidos. Este recurso que suele estar desaconsejado por las academias e instituciones es, no obstante, muy utilizado en este sector profesional, porque permite perpetuar para la posterioridad a un investigador o médico, darle popularidad y reconocimiento. Los términos eponímicos, pero, son unidades semánticamente muy opacas puesto que el nombre propio no nos dice nada del referente a no ser que conozcas el contexto y la vida de la persona. Este fenómeno motiva que surjan en paralelo términos equivalentes más descriptivos con patrones morfosintácticos propios de los términos médicos canónicos (*enfermedad de Lowe/distrofia oculocerebrorenal, síndrome de Allan-Herndon-Duley/déficit del transportador de monocarboxilato 8, síndrome de Crandall/síndrome de apolecia-sordera-hipoginadismo, enfermedad de Abrami/ictericia hemolítica, enfermedad de Albert/bursitis retrocalcania, enfermedad de Ballingall/micetoma, enfermedad de Bazin/psoriasis bucal, enfermedad de Crocq/acrocianosis*, etc.).

Otra de las características de los textos médicos y que supone uno de los fenómenos actualmente más preocupantes es la abreviación de unidades terminológicas poliléxicas, es decir compuestas por más de una palabra. Las siglas, acrónimos y abreviaturas son cada vez más abundantes en los documentos médicos. Este fenómeno está relacionado con la precisión y con la economía del lenguaje. Por un lado, se tiende a ser muy preciso con lo que el hablante crea y utiliza unidades poliléxicas; estas unidades que suelen ser descriptivas a su vez son demasiado largas. La longitud excesiva de estas unidades propicia que el hablante las acorta por una ley universal de economía lingüística, lo que las hace más manejables, pero también las vuelve totalmente opacas semánticamente (*acidosis tubular distal/ATD, reacción en cadena de la polimerasa/parada cardiorrespiratoria/proteína C reactiva/reflejo cutáneo plantar/PCR; tratamiento antirretroviral de alta eficacia/TAAE*).

En síntesis, los términos médicos se pretende que sean unívocos, inequívocos e inocentes, pero cuando se usan en discursos, pueden ser semánticamente obscuros, polisémicos, poco precisos e incluso connotados peyorativamente (ESTOPÀ, 2013).

Desde el punto de vista gramatical, los términos en los diccionarios o bases de datos suelen ser siempre nombres. En cambio, desde la óptica de la Teoría Comunicativa de la Terminología de la Escuela de Barcelona (CABRÉ, 1999; ESTOPÀ, 2001; CABRÉ; DOMÈNECH; ESTOPÀ, 2018), de base lingüística y que basa sus observaciones en textos reales, las unidades que transmiten significado especializado son de diferentes categorías, aunque es cierto que los nombres son las más prototípicas. Así, los términos nominales conviven con verbos (*inocular, desinfectar, operar, inyectar*), adjetivos (*renal, ocular, clínico, hospitalario), adverbios (por vía oral, secundariamente, clínicamente*) que vehiculan significado especializado relevante en el ámbito médico. También se usan unidades de otros sistemas no lingüísticos, como los símbolos de la química (O, N, C) o los nombres en latín de nomenclaturas científicas (*Rattus rattus, femur femoris, rotula*). En los textos también encontramos unidades que superan los límites de las unidades léxicas y que configuran la fraseología de este ámbito del saber (*radiografía simple de brazo derecho, TAC de abdomen, inyección de ATB, lavado de estómago en intoxicación autolítica, colocación de un yeso en un pie, artrocentesis de la rodilla*). En consecuencia, es muy importante estudiar la terminología en contexto, en uso, dentro de un texto, porque se tiene una aproximación real que se distorsiona si se estudia solo a partir de un diccionario o vocabulario.

3 Los informe médicos

Son muchos los tipos textuales que vehiculan conocimiento médico: artículos, informes, manuales, tesis doctorales, libros de divulgación, clases magistrales, conferencias, ponencias, etc. De entre todos hay un tipo de texto especial y muy relevante para la práctica médica en centros sanitarios: el informe médico.

El informe médico que forma parte de la historia clínica de un paciente es el documento donde se refleja el proceso diagnóstico o terapéutico realizado durante una asistencia sanitaria. Forma parte de la historia clínica. Es un texto que está escrito por un médico, a veces incluso por más de uno, aunque podemos decir que el emisor es siempre homogéneo: es un especialista en el ámbito con una formación amplia y académica en la materia. En esto el informe médico no difiere de otros tipos de documentos especializados. Lo que sí que es especial es el poco tiempo en el que el médico redacta este documento y a veces lo redacta en períodos discontinuos temporalmente.

Pero lo que es realmente complejo de un informe médico es el destinatario. Así, desde el punto de vista comunicativo, el informe médico es un tipo de texto especializado que presenta una complejidad elevada, puesto que suele estar escrito por un único emisor especialista, el médico (aunque a veces puede estar escrito por diversos médicos o profesionales de la sanidad), pero se dirige a destinatarios diversos, tanto desde el punto de vista cognitivo – nivel de conocimiento del ámbito temático de la medicina – como desde el punto de vista emocional – distintas inquietudes en relación con el texto.

De manera que, aunque en primera instancia y por ley el informe médico se dirige al paciente o a su familia, también son sus destinatarios el médico especialista o generalista, otros profesionales de la salud, documentalistas, investigadores, juristas y gestores sanitarios. Podemos distinguir tres grandes tipos de posibles destinatarios del informe médico:

a) especialistas para los que el informe médico cumple una función informativa para la asistencia. Se trata de otros médicos que continúan la asistencia del paciente en el mismo centro o en otros centros sanitarios: una importante proporción de enfermos no están curados cuando reciben el alta y requieren diversos tratamientos y un seguimiento de su enfermedad. Por ello, el informe médico debe resumir el proceso que ha seguido el paciente y la situación en la que se encuentra de la forma más simple, completa y eficiente permitiendo una mejor gestión de los recursos y unos cuidados adecuados.

b) codificadores o personal sanitario que asigna los códigos correspondientes a los diagnósticos y procesos asistenciales, a través de la información que les ofrece el informe médico. Esos códigos serán utilizados para la recuperación de la información con finalidades que pueden ser económicas, de control de la calidad o de investigación.
c) pacientes para los que el informe médico debería cubrir una función informativa: la comprensión del diagnóstico y el tratamiento a seguir después de salir de un centro hospitalario deberían ser prioridad primera.

Además de estos destinatarios principales, cabe mencionar también otros posibles destinatarios del informe médico, como los gestores sanitarios, que persiguen una práctica profesional eficiente que garantice una atención sanitaria equitativa y adecuada, y los juristas, que intervienen en los cada vez más numerosos casos con de denuncias sanitarias. El informe médico, por lo tanto, es un texto dirigido a destinatarios con unas capacidades lingüísticas y cognitivas muy diversas, y que, por lo tanto, debe cubrir funciones muy distintas: legal, científica, biográfica, comunicativa, evaluativa (DELÀS, 2005). Esta diversidad es una característica muy significativa del destinatario del informe médico que explica, en gran parte, su elevada complejidad, porque es extremadamente difícil escribir un texto que pueda satisfacer las necesidades de personas con un marco mental tan distinto y con inquietudes tan diferentes sobre el texto. Autores como Safeer y Keenaan (2005) insisten en la idea de que el tipo de redacción y el nivel de comprensión lectora que exigen los informes médicos provocan una brecha comunicativa importante que acrecienta una desigualdad entre pacientes y médicos.

Partimos de una situación real: el paciente recibe un informe de un médico especialista y en la mayoría de los casos no lo entiende, tiene problemas importantes para entenderlo o tiene diversas dudas. Lo que suele producir estrés, angustia, malas interpretaciones que comportan mala gestión de su enfermedad. Muchas veces acude al médico de cabecera para que se lo interprete, le explique qué tiene y qué debe hacer. Esta situación que se repite a diario encarece el sistema sanitario y lo complica logísticamente. Los pacientes que no acuden a su médico,

que suelen ser los más hábiles tecnológicamente, acuden a Dr. *Google* para buscar información que les permita entender su informe. Pero muchas veces ocurre el efecto contrario: falsas interpretaciones, incomunicación, sobrecarga cognitiva, etc. Los informes no son fáciles de leer por muchos motivos, pero la lengua no debería ser un obstáculo y ahora por ahora lo suele ser.

4 ¿Cómo son lingüísticamente los informes médicos?

Podemos acercarnos a los informes y analizar la información cognitiva como han hecho algunos investigadores (CONTHE *et al*., 2010), que han constatado que un informe médico de alta bien planificado evitaría un tercio de reingresos. Pero también podemos analizarlos des del punto de vista lingüístico y discursivo. Lingüísticamente, los informes médicos presentan muchos problemas que se convierten en retos.

Los problemas que acarrea la comunicación médico-paciente y en concreto la calidad lingüística de la información escrita que recibe el paciente como los informes médicos es el objetivo central del proyecto JUNTS-RECERCAIXA (https://www.upf.edu/web/medicina_comunicacio). En este proyecto, la monitorización de la comunicación médico-paciente se ha abordado de maneras distintas:

a) encuestas psicolingüísticas de profundidad con 15 familias;
b) análisis de 50 informes con programas automáticos de legibilidad (PORRAS-GARZÓN; ESTOPÀ, 2018);
c) estudio microlingüístico de 50 informes a través de parámetros lingüísticos que dificultan la comprensión;
d) encuestas de comprensión de un informe real y de un informe mejorado y enriquecido (ESTOPÀ; DOMÈNECH-BAGARIA, 2018).

Los resultados de estos trabajos se han divulgado en ponencias y publicado en artículos que pueden consultarse en la página web del proyecto. En el presente artículo nos interesa mostrar algunos de los resultados del estudio microlingüístico que llevamos a cabo con un corpus de 50 informes médicos reales. En primer lugar establecimos una batería de elementos lingüísticos que obstaculizan la

comprensión. Se trata de problemas de distinta naturaleza lingüística que hemos agrupado en dificultades de tipo:

a) pragmático-semántico;
b) sintáctico;
c) léxico;
d) ortotipográfico.

Cada uno de estos grandes tipos se ha desglosado en diversos parámetros concretos. Por ejemplo, dentro del léxico se incluyen el número de siglas utilizadas, el número de términos con formantes grecolatinos, el número de términos en relación con las palabras del texto, etc. O en el caso del grupo de dificultades sintácticas se incluyen parámetros relativos al número de oraciones, al número de oraciones sin verbo explícito, al tipo de verbos utilizado, etc.

En general, podemos avanzar que los informes médicos analizados son textos expositivos que se caracterizan por un estilo nominal muy marcado, con muy pocos verbos, que con frecuencia son impersonales o se usa la voz pasiva. Léxicamente, presentan una elevada densidad terminológica, un uso abusivo de siglas y símbolos sin expandir, y una elevada presencia de términos semánticamente opacos. Ortotipográficamente, contienen abundantes errores ortográficos y tipográficos. Sintácticamente muestran una puntuación con frecuencia inadecuada y una falta de conectores entre párrafos y entre oraciones. Semánticamente carecen de metáforas explicativas, analogías y paráfrasis que favorezcan la comprensión de expresiones terminológicas opacas. Pragmáticamente, suelen ser documentos despersonalizados, neutrales, en lo que importa es la enfermedad y su tratamiento, mientras que los sentimientos del paciente son poco relevantes. En consecuencia, con todos estos déficits lingüísticos es fácil deducir que el paciente tiene dificultades para entender su informe médico.

En esta parte final de este artículo profundizaremos más algunos de los aspectos terminológicos.

5 Problemas terminológicos

Algunas encuestas realizadas a pacientes (ESTOPÀ; DOMÈNECH, 2018) y también a médicos (NAVARRO, 2016) constatan que la terminología es uno de los principales obstáculos de comprensión para el destinatario de un informe médico. Los términos asustan, los términos no se entienden, los términos suscitan imaginarios negativos, los términos cuestan de pronunciar, los términos colapsan, se malinterpretan, etc. Para muchos receptores, en especial para los pacientes, la presencia de términos quizás sea uno de los principales problemas para entender los informes médicos, su característica más prominente y también la más compleja. ¿Pero qué parámetros son los que hacen que la terminología sea un problema si son palabras de la lengua?

Los obstáculos de comprensión que se derivan del uso de terminología se pueden resumir, siguiendo a Estopà i Montané (2018, en prensa), en los siguientes cinco parámetros:

a) concentración de terminología: el número de términos que contienen los informes médicos es muy elevado en relación con el número de palabras del texto;
b) opacidad semántica: los términos no suelen resultar semánticamente transparentes, sino que causan problemas de comprensión a sus destinatarios (especialmente a los pacientes y familiares);
c) confusión semántica: los términos, como a menudo resultan difíciles de comprender, pueden llevar a equívocos en relación con su significado;
d) ambigüedad semántica: los términos presentan variación y polisemia, lo que puede provocar que se entiendan de varios modos o que conduzcan a distintas interpretaciones y, por consiguiente, que generen dudas e incertidumbre.

Por ejemplo, en los informes médicos se da una concentración muy alta de términos: se calcula que se utiliza una media de 79 términos por texto, en textos que tienen de media unas 400 palabras. Así pues, se estima que el porcentaje de términos por informe es del 22,6 %. Esta ratio resulta muy elevada, de acuerdo con el estudio de corpus realizado por Domènech-Bagaria (2007), en

que se analizó la densidad terminológica ("medida como el número de términos detectados en relación con el número de palabras de cada texto") de varios textos con niveles de especialización distintos, el cual observó que el porcentaje de densidad terminológica se sitúa entre el 9 % y el 21,5 % en textos jurídicos, y entre el 10 % y el 18,6 % en textos sobre genómica.

En el ejemplo siguiente se puede observar como la concentración de términos en los informes médicos es muy alta. Y llega a ser del 100% en apartado como los antecedentes o sobre todo el diagnóstico:

Antecedentes personales:
- Síndrome de Lowe (oculocerebrorenal) con tubulopatía compleja en control por Nefrología, difícil de tratar, con acidosis metabólica, proteinuria, hipercalciuria y nefrocalcinosiso y raquitismo resistente a vitamina D secundário;
- Crisis convulsivas: tratamiento farmacológico y control por Neurología

Diagnóstico:
- Síndrome de Lowe. Tubulopatía. Nefrocalcinosis
- Epilepsia sintomática.

Otro apartado en las que el control y el cuidado de la terminología deberían ser extremos es el del tratamiento o recomendaciones al alta. Las consecuencias más directamente asociadas con la salud del paciente son las relacionadas con los medicamentos y sus dosis. Por eso, por citar un aspecto preocupante, el uso excesivo de abreviaciones cuando se redacta este apartado del informe médico puede representar un problema de salud pública pues conlleva errores en la medicación que afectan directamente a la salud del paciente. Veamos un ejemplo:

Tratamiento al alta
- Rocaltrol 0.25 mcg caps: 1/mañana, excepto los domingos.
- Ibercal 500 mg/5 ml, sol: 5 ml/12h.
- Solución Sohl 15 ml/8h.
- Valproato (Depakine 200mg/ml): 0,9 ml cada 8 horas.

Las abreviaciones, la poca sistematicidad y la nominalización de los enunciados provoca una comprensión baja o distorsionada del texto. Esto no es una excepción sino que se repite en muchos informes:

Tratamiento

- TORASEMIDA 10mg, 2 comp/d 1 - 1 - 0
- ESPIRONOLACTONA 25 mg, 1 comp/d 1 - 0 - 0
- CARVEDILOL 6.25 mg, 1/2 co/12h 1/2 - 0 - 1/2
- SINTROM UNO 1mg, según pauta 0 - 0 - x
- OMEPRAZOL, 20mg, 1 comp/d 0 - 0 - 1
- FEROGRADUMET, 1 comp/d 1 - 0 - 0
- RISPERDAL solución, 0.5 ml por la noche 0 - 0 -1/2
- Dieta con poca sal. Ingesta abundante de líquidos. Actividad física según posibilidades.
- Seguimiento por el Equipo de Atención Primaria y Cardiología CAP.

5 Conclusiones

La comunicación, privilegio de la condición humana, parece estar "enferma" en la práctica médica, en la comunicación médico-paciente y tampoco es clara entre médico-médico. Pero por suerte en los últimos años hay voces y nuevas directrices que contemplan los derechos de los enfermos que han vuelto a centrar en la persona (la medicina centrada en el paciente), pues han visto que había ganancias muy diversas para todos. En el futuro los enfermos tendrán un papel central en la toma de decisiones sobre su salud.

En este nuevo giro las **habilidades de comunicación** son extremadamente relevantes para mejorar la relación médico-enfermo, en general, y de manera especial en las enfermedades graves y en las enfermedades crónicas. De aquí la importancia que el médico tenga formación en comunicación y en redacción adecuada de textos destinados a los pacientes. El médico debería concienciarse del valor que puede tener para las personas, y también para el sistema sanitario, **entender** sus informes médicos: entender qué les pasa, por qué y qué tienen que hacer.

En conclusión, los informes médicos (los textos clínicos en general) presentan peculiaridades muy especiales. Conocer sus características, es decir contar con análisis microexhaustivos y desde diferentes puntos de vista de la calidad de la información (no del tipo de datos), nos permite a) adecuar los documentos a las necesidades reales; b) diseñar herramientas de asistencia para generar documentos más sistemáticos, más comprensibles; c) concienciar a la comunidad emisora que la receptora debe entenderle. Solo siendo capaz de generar estas **empatías** entre los profesionales de la salud y los pacientes podemos empezar a pensar en una sanidad más inclusiva.

Referencias

ANDERSON, D. **Medical terminology**. 11. ed. S.l. Medical Creations, 2016.

BERNABEU, J. *et al*. **El llenguatge de les ciències de la salut**. València: Generalitat de València, Conselleria de Sanitat i Consum, 1995.

BORRELL, F. **Entrevista clínica**: manual de estrategias prácticas. Barcelona: Semfyc ediciones, 2004.

CABRÉ, M. T. **La terminologia, representación y comunicación**: una teoría de base comunicativa y otros artículos. Barcelona: Institut Universitari de Lingüística Aplicada, 1999.

CABRÉ, M. T.; DOMÈNECH-BAGARIA, O.; ESTOPÀ, R. B. **La terminologia avui**: termes, textos i aplicacions. Barcelona: Editorial UOC, 2018.

CONTHE, P. *et al*. Consenso para la elaboración del informe de alta hospitalaria en especialidades médicas. **Medicina clínica**, v. 134, n. 11, p. 505-510, 2010.

DELÀS, J. **Informes clínics, eines de comunicació**. Barcelona: Quaderns de la bona praxis Collegi Oficial de Metges, 2005.

DOMÈNECH-BAGARIA, O. La noció de text especialitzat des de la perspectiva de la teoria comunicativa de la terminologia. *In*: LORENTE, M. *et al*. **Estudis de lingüística i de lingüística aplicada en honor de M. Teresa Cabré Castellví**. Barcelona: Universitat Pompeu Fabra, Insitut Universitari de Lingüística Aplicada, 2007. p. 241-253.

ESTOPÀ, R. B. El léxico de la salud: necesidades terminológicas y acceso al conocimiento especializado. *In*: BENTES PINTO, V.; CAMPOS, H. de H. **Diálogos paradigmaticos**: informacação para a àrea de saúde. Fortaleza: Editora UFC, 2013.

ESTOPÀ, R. B.; DOMÈNECH-BAGARIA, O. Diagnóstico del nivel de comprensión de informes médicos dirigidos a pacientes y familias afectados por una enfermedad rara. *In*: CONGRESO INTERNACIONAL DE AESLA, 36., 2018, Cádiz. **Anais** [...]. Cádiz: [s.n.], 2018.

ESTOPÀ, R. B.; MONTANÉ, A. La terminología: principal obstáculo de incomprensión de los informes medicos. *In*: ESTOPÀ, R. B. **La comprensión de los informes médicos**. 2018.

ESTOPÀ, R. Les unités de signification spécialisées élargissant l'objet du travail en terminologie. **Terminology**, v. 7, p. 217-237, 2001.

FALCÓN, M.; LUNA, A. Alfabetización en salud: concepto y dimensiones: proyecto europeo de alfabetización en salud. **Rev. Comunicación y Salud**, v. 2, n. 2, p. 91-98, 2012. Disponible a: http://revistadecomunicacionysalud.org/index.php/rcys/article/view/35. Accedir a: 15 maio 2018.

GALARZA, C. A. *et al*. Calidad de comunicación para una mejor salud. Oportunidad y desafío de la world wide web. **Revista Nexo**, v. 23, n. 1, 2003.

GARCÍA PALACIOS, J. El lenguaje médico, algo más que información. **Panacea**, v. 5, n. 16, p. 135-140, jun. 2004. Disponible a: http://www.

medtrad.org/panacea/IndiceGeneral/n16_tribuna_GarciaPalacios.pdf. Accedir a: 15 maio 2018.

GUTIÉRREZ, B. **La ciencia empieza en la palavra**: análisis e historia del lenguaje científico. Barcelona: Península, 1998.

KURTZ, Suzanne M. Doctor-patient communication: principles and practices. **Canadian Journal of Neurological Sciences**, v. 29, supl. 2, S23-9: S23-9, jun. 2002.

MOORE, P. *et al*. La comunicación médico-paciente: ¿Cuáles son las habilidades efectivas? **Rev Med Chile**, v. 138, n. 8, p. 1047-1054, ago. 2010.

NAVARRO, F. El inglés, idioma internacional de la medicina. Causas y consecuencias de un fenómeno actual. **Panacea**, v.2, n. 3, p. 35-51, 2001.

NAVARRO, F. La precisión del lenguaje en la redacción médica. **Cuadernos de la Fundación del Dr. Antonio Esteve**, v. 17, p. 89-104, 2016.

PIÑERO, J. M. L.; TERRADA, M.L. **Introducción a la terminología médica**. Barcelona: Manuales Salvat, 1990.

PORRAS-GARZÓN, J. M.; ESTOPÀ, R. B. Recursos para hacer análisis de comprensión a textos médicos escritos: análisis cuantitativo de tres casos médicos. In: JORNADA REALITER, 13., 2018, Barcelona. **Anales de congreso** [...]. Barelona: UPC, 2018.

SAFEER, R. S.; KEENAN, J. Health literacy: the gap between physicians and patients. **Am. Fam. Physician**, v. 72, n. 3, p. 463-468, ago. 2005. Disponible a: http://www.ncbi.nlm.nih.gov/pubmed/16100861. Accedir a: 15 mai. 2018.

SHUKLA, A. K.; YADAV, V. S.; KASTURY, N. Doctor-patient communication: an important but often ignored aspect in clinical medicine. **JIACM**, v. 11, n. 3, p. 208-211, 2010.

SØRENSEN, K. et al. Health literacy in Europe: comparative results of the European health literacy survey (HLS-EU). **European Journal of Public Health**, v. 25, p. 1053-1058, 2015.

TUSON, J. **Una imatge no val més que mil paraules**. Barcelona: Grup62, 2002.

O Facebook como espaço tecnológico e discursivo d@s usuári@s

Daniel Thiago Freire da Silva
Izabel França de Lima
Maria da Luz Olegário

1 Introdução

As transformações pelas quais o mundo vem passando têm modificado, dentre outros aspectos, a maneira como as pessoas se relacionam umas com as outras. A inserção do computador conectado à internet e, mais recentemente, de outros aparatos tecnológicos, no cotidiano social, como os dispositivos móveis, reformulou a interação humana, sobretudo quando se percebe que as barreiras sociais, culturais, espaciais e temporais foram diluídas e já não se apresentam como empecilhos para a construção e manutenção de redes relacionais entre os indivíduos.

Martino (2015) alerta que o ciberespaço também se tornou o ambiente adequado para que as pessoas igualmente transpusessem suas emoções negativas e suas frustações. Para ele, porém, esse espaço não é o causador desses sentimentos hostis, mas sua arquitetura aberta, associada à sensação de que todas as coisas se criam e se desfazem ao passo da instantaneidade, possibilita o compartilhamento e a propagação de mazelas sociais que, agora, são tão presentes no ciberespaço, tal como na sociedade desconectada.

É no ciberespaço que ocorre uma representação simbólica da vida material dos humanos. Essa perspectiva enfatiza a ideia de que a humanidade sempre viveu a articulação desses dois institutos, já que para perceber e vivenciar a materialidade da vida (realidade concreta) se produz e se consome sinais e representações diversas (realidade virtual), conforme Castells (2016). Assim, o virtual e o real não se opõem, estando essas entidades interligadas em um processo de mútua influenciação e sinergia (LEVY, 2010).

As possibilidades disponíveis no ciberespaço se puseram como alicerce para o surgimento das comunidades virtuais: coletivo de pessoas que compartilha, entre si, discursos que evidenciam seus interesses e objetivos; que se beneficia das facilidades desse ambiente sem território para a construção e manutenção de relacionamentos. Formam-se redes no ciberespaço. Uma teia de nós que deu luz ao que conhecemos, hoje, como redes sociais virtuais ou apenas redes sociais: fenômeno relativamente novo e que ganhou grande notoriedade na sociedade atual.

O Facebook surge no ano de 2004 despontando com a proposta inicial de possibilitar a interação entre um pequeno grupo de estudantes da Universidade de Harvard, nos Estados Unidos (AMORIM, 2011).

Hoje, diversas são as ferramentas e utilidades ofertadas por essa rede social e, dentre elas, está a possibilidade de manter-se informado/a devido ao compartilhamento de notícias. Diante desse oceano de informações, muitos desses sujeitos veem, nesse ambiente, uma oportunidade para expressar suas opiniões e pontos de vista acerca das mais variadas temáticas circulantes na rede, ainda que muito do que seja exposto ali fuja das práticas de respeito e tolerância que se esperam da humanidade. Amparados/as por um suposto anonimato e por uma liberdade de expressão ilimitada, esses sujeitos propagam manifestações de preconceito e discriminação nesses espaços virtualizados (PENA, 2017).

O surgimento do ciberespaço e, mais recentemente, das redes sociais, em especial o Facebook, possibilitaram que os sujeitos construíssem, nesses espaços, vontades de verdade forjadas por uma suposta liberdade de expressão ilimitada. Nessa perspectiva, objetivamos compreender o discurso d@s usuári@s no Facebook, a partir da Análise de Discurso.

2 Das comunidades virtuais às redes sociais virtuais

A possibilidade de se conectar com pessoas de todos os lugares do mundo culminou com o surgimento das comunidades virtuais e, consequentemente, o estabelecimento de novos laços interpessoais. Assim como ocorre na vida desconectada, esse agrupamento virtual tece sua estrutura a partir de interesses

e afinidades comuns, em um processo de cooperação ou de troca de informações virtualizadas.

A palavra comunidade já denota essa afinidade mútua, visto que a raiz desse termo "comum" é designada àquilo que pode ser dividido, compartilhado, assim como ocorre na palavra comunicação, por exemplo. Podemos, então, entender o fenômeno das comunidades virtuais como um conjunto de relações interpessoais no ciberespaço, comunidades essas onde os sujeitos compartilham assuntos afins por determinado período de tempo. O desenvolvimento das comunidades virtuais acompanha, em geral, contatos e interações de todos os tipos. Uma comunidade virtual não é irreal, ilegítima, imaginária ou ilusória, trata-se simplesmente de um coletivo volátil que se organiza por meio das facilidades disponibilizadas pelo ciberespaço (LEVY, 2010; MARTINO, 2015).

Assim, as comunidades virtuais podem servir, também, como um palco para o qual os sujeitos transportam todos seus sentimentos e frustrações, tornando-as, desse modo, em um local propício para o desenvolvimento de problemas que já eclodem na sociedade. Não é o ciberespaço que constrói o discurso inseminado, mas a arquitetura desse espaço, em especial no que se refere às comunidades virtuais, pode ser utilizada como dispositivo para compartilhamento de hostilidades, potencializando e transmutando mazelas sociais, sobretudo quando há conexão com outros que pensam da mesma forma (MARTINO, 2015).

Isso reafirma a impossibilidade de dissociação entre o que se vive no ciberespaço e o que se passa na vida concreta, o que nos faz ir além e imaginar que essas comunidades virtuais abertas no ciberespaço são um prelúdio daquilo que atualmente conhecemos como redes sociais virtuais, reconhecidamente externadas nos discursos.

Hoje muito se fala sobre redes sociais. Usamos esses termos, quase sempre, para designar sites ou aplicativos atuais que, inseridos no ciberespaço, facilitam o encontro, o relacionamento entre as pessoas e o compartilhamento de conteúdos. Contudo a humanidade, desde sua existência, cria e mantém vínculos relacionais entre seus pares.

Podemos entender as redes como um conjunto de pontos interconectados que estabelecem comunicação entre si em uma estrutura complexa de comunicação,

dentro da qual todos esses nós mantêm interação em múltiplas e simultâneas ligações. As redes são estruturas abertas e que estão em constante movimento. Em outras palavras: redes, sobretudo inseridas no contexto da era da informação, podem ser vislumbradas como um conjunto de nós intercalados, dentro dos quais as conexões assumem o lugar do que antes seria apenas linearidade. A rede é, portanto, uma estrutura hipertextual, não-linear, descentralizada, flexível, dinâmica e mutável (CASTELLS, 2016; MARTINO, 2015).

A denominação 'redes sociais' passou a ser utilizada com mais frequência para designar um fenômeno novo e de grande força na sociedade atual. Esse termo é direcionado para descrever uma estrutura composta por pessoas ou organizações conectadas por um ou vários tipos de relações que compartilham informações, valores e objetivos comuns, sendo, dessa forma, um conjunto de participantes que unem ideias e recursos em torno de valores e interesses compartilhados (TOMAÉL; ALCARÁ; DI CHIARA, 2005; VICENTE, 2014).

Para Martino (2015), as redes sociais inseridas nos espaços virtuais podem ser concebidas como uma estrutura flexível e dinâmica, na qual os seres humanos se relacionam sem que exista, com isso, uma estrutura hierárquica rígida. A partir dessa definição, portanto, podemos extrair duas características essenciais das redes sociais virtuais: a dinâmica e a flexibilidade.

A primeira se direciona a maneira com a qual os participantes interagem nas redes sociais, sendo que cada rede social tem sua própria dinâmica. Quanto à flexibilidade, ela se refere à maleabilidade dos vínculos ali criados. O estabelecimento, a manutenção e a diluição dessas conexões se dão a qualquer instante, distintamente do que ocorre em grande parte das relações humanas geradas no mundo concreto, em que as redes sociais tomam, por alicerces, valores e obrigações compartilhados.

Estruturalmente, as redes sociais são compostas por elementos (atores/atrizes) interligados por teias ou nós. É importante ressaltar, principalmente por estarmos tratando desses agrupamentos dentro do ciberespaço, que o papel de ator/atriz não precisa ser, necessariamente, desempenhado por pessoas. As relações, além de se darem entre os indivíduos inseridos no ciberespaço, ocorrem também de forma intrarrelacional, ou seja, a maneira como esses sujeitos estabelecem

relações entre si vai reverberar em outras conexões, em outras relações nesse espaço, isto é, "uma perspectiva mútua e recíproca sobre a maneira como as pessoas interagem" (MARTINO, 2015, p. 57).

Apesar de atualmente as redes sociais terem se tornado um espaço de grande visibilidade na sociedade, é necessário apontar que estas sempre existiram, o que há de novidade, porém, é o crescimento das redes sociais no ciberespaço, estimulado pelas facilidades e disseminação do uso da internet.

3 Das redes sociais, o Facebook

Em 28 de outubro de 2003, o então estudante Mark Zuckerberg lançou o site *Facemash*, que é considerado o precursor daquilo que hoje conhecemos como Facebook[10]. O surgimento do Facebook, tal como ele é atualmente, data precisamente de fevereiro de 2004 e teve sua proposta inicial direcionada para o uso exclusivo de estudantes da Universidade de Harvard – Estados Unidos, sendo utilizada como uma ferramenta básica para comunicação entre esses indivíduos (AMORIM, 2011). Em maio do mesmo ano, essa rede social passou a permitir o acesso de discentes de outras universidades e foi somente em dezembro de 2004 que o Facebook permitiu o acesso indiscriminado de novos/as usuários/as. Hoje, através desse site, o/a interessado/a pode criar um perfil virtual e, a partir de então, estabelecer uma rede de contatos e compartilhamento de conteúdos, fotos, vídeos, notícias, textos, dados pessoais com outras pessoas em qualquer lugar do planeta.

> O Facebook é um site de redes sociais utilizado em larga escala e mundialmente. Além disso, caracteriza-se por, no passar dos anos, ter ampliado suas possibilidades de participação inserindo novas funcionalidades, principalmente inspiradas pela própria utilização da plataforma por seus usuários. Seu objetivo parece ser o de abranger todas as formas de compartilhamento e leitura de conteúdo pessoal ou informativo

10 Inicialmente o Facebook recebeu o nome de The Facebook. Esse primeiro nome remete aos livros de fotos e dados bibliográficos tipicamente distribuídos aos(às) estudantes de Harvard (AMORIM, 2011).

> online. Tornando-se, assim, um ambiente complexo onde se torna possível observar uma grande variedade de dados referentes à formação de identidade e trocas de conteúdos com os quais os sujeitos se identifiquem. (MAZZOCATO, 2014, p. 23).

Desde que se tornou empresa pública com vendas de ações na bolsa de valores, em 2012, o Facebook vem apresentando ganhos financeiros acima do esperado, ainda que, em seus primeiros momentos, nesse ramo de negócio, tenha sido frágil (OLHAR DIGTAL, 2014). Ao revelar seu relatório financeiro de 2016, o Facebook noticiou lucros acima de cinco bilhões de dólares. Quando se compara ao ano anterior, cujos números alcançaram a marca de três bilhões e meio de dólares, percebemos um considerável crescimento. (CANAL TECH, 2016).

Já no ano de 2017, o Facebook alcançou a marca de mais de dois bilhões de usuários (GLOBO, 2017) e é considerado, portanto, a maior e mais usada rede social do mundo, sendo, inclusive, também, a mais popular entre o público jovem mundial. Apesar das acentuadas marcas da exclusão social que perpassam a sociedade brasileira, o Brasil ocupa, hoje, a segunda posição no mundo em quantitativo de contas de usuários/as ativos/as, com mais de 71 milhões de usuários cadastrados, ficando atrás, apenas, dos Estados Unidos (GONDIM, 2017).

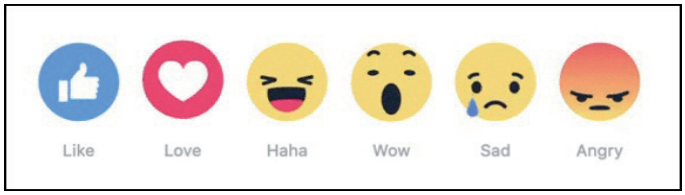
Dentre as funcionalidades do Facebook, podemos destacar os recursos que permitem a criação e divulgação de eventos sociais, virtuais ou não; a criação e participação de grupos de interesses e discussões específicos e a possibilidade de compartilhar, em tempo real, as atividades desenvolvidas por seus/as usuári@s, como, por exemplo, que livros estão lendo, quais músicas escutam, programa de tv que estão assistindo, onde e com quem viajam e, até mesmo, o que estão comendo ou bebendo.

Mais recentemente, em uma de suas atualizações, o Facebook lançou a ferramenta *"Reactions"* (Reações), que permite aos usuários/as que estes demonstrem, através dessas representações virtuais, como reagem a determinado

conteúdo exposto nessa rede social. Cada *reaction* representa virtualmente um sentimento distinto sobre determinado tema (Figura 1).

Sendo assim, quando o sujeito usa o botão *Like*/Curtir para alguma postagem, por exemplo, isso denota uma reação positiva ao conteúdo que lhe foi apresentado. Com o botão *Love*/Amei, podemos inferir que houve uma sensação além do simples gostar, que foi algo mais interessante e/ou de maior identificação. A reação *Haha* é direcionada para expressar que algo teve um tom cômico, divertido ou irônico. Já o botão *Wow*/Uau denota surpresa ou espanto. *Sad*/Triste é usado para demostrar tristeza sobre algum assunto e *Angry*/Grr representa raiva e desaprovação a determinado conteúdo compartilhado.

Figura 1 - Função Reações no Facebook



Fonte: TechTudo (2016).[11]

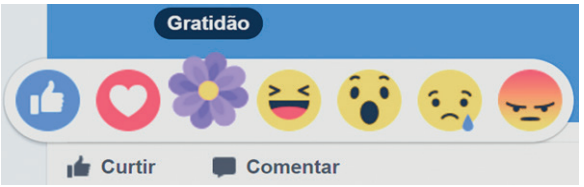
Essa novidade foi uma maneira encontrada pelo Facebook para tornar a rede social ainda mais próxima da vida que acontece do lado de fora da tela do computador.

> [...] a mudança foi feita porque os usuários sentiam que deveriam existir outras formas de "expressar de forma rápida e fácil como algo que você vê no *News Feed* te emocionou de alguma forma". Quando alguém usa o Facebook para anunciar o falecimento de um ente querido, por exemplo, o botão de "Curtir" não é a melhor forma de reagir; agora, será possível usar a reação "triste", por exemplo (TECNOBLOG, 2016).

11 Disponível em: https://www.techtudo.com.br/noticias/noticia/2016/02/facebook-libera-novos-botoes-de-reacao-alem-do-curtir-no-mundo-todo.html

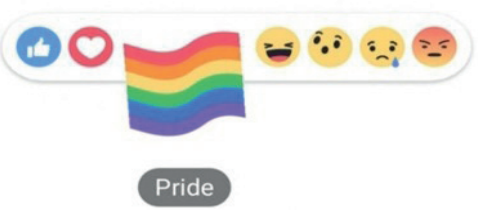
Eventualmente, o Facebook disponibiliza a seus/suas usuários/as, de forma temporária, outras *Reactions*/Reações. Foi assim com a reação Gratidão, representada por uma flor margarida lilás, direcionada para a comemoração do dia das mães (Figura 2), e com a reação temática *Pride*/Orgulho, representada pela bandeira do arco-íris (Figura 3), como uma campanha de apoio e homenagem ao mês do orgulho LGBTI, sigla para lésbicas, gays, bissexuais, travestis, transexuais, pessoas trans e intersex.

Figura 2 – Reação Gratidão em comemoração ao dia das mães



Fonte: Extraído de facebook.com (2017).

Figura 3 – Reação Orgulho em comemoração ao dia do Orgulho LGBTI



Fonte: Extraído de facebook.com (2017).

Essas possibilidades representaram uma nova maneira de interação nesta rede social virtual, agregando sentidos nas formas como os/as usuários/as do Facebook se posicionam sobre os conteúdos que surgem na tela.

Figura 4 – Página Inicial do Facebook

Fonte: Extraído de facebook.com (2017).

O Facebook, cuja página incial é mostrada na figura 4, reformulou a maneira como os indivíduos estabelecem relacionamentos, dividem conteúdos e opiniões, com a criação de novas ferramentas digitais, transformou, também, o modo pelo qual seus/suas usuários/as produzem verdades forjadas em seus discursos.

Assim, com diversas inovações e possibilidades, o Facebook, paulatinamente, torna-se uma ferramenta indissociável do convívio das pessoas ao redor do mundo e, com isso, é pertinente considerar que essa rede social virtual, em especial, firma-se, na verdade, como um espelho que reflete, por entre as teias do ciberespaço, as características, potencialidades e fraquezas da sociedade off-line contemporânea.

4 O Facebook como espaço discursivo

O ciberespaço provocou e provoca mudanças de paradigmas nos comportamentos sociais, além de colocar em cena novas formas de se expressar. Nesse contexto, o espaço virtualizado desponta como local oportuno para manifestações de posicionamentos, ainda que muitas delas tragam em seu arcabouço expressões que se afastam da urbanidade e do respeito coletivo.

Em tempos de redes sociais, podemos observar que os discursos produzidos pelos sujeitos se dão por meio de novas linguagens e novos símbolos. Quando

afunilamos mais esse universo e tornamos nosso foco para o *Facebook*, podemos perceber ainda que, além dos textos produzidos por palavras, os sujeitos produzem seus discursos por meio de curtidas em *posts*, compartilhamentos de conteúdos, *emojis*[12], *memes*[13], *hashtags*[14], figuras, vídeos.

A esse respeito, trazemos o pensamento de Orlandi (1999), quando afirma que os sentidos do discurso se expandem para além do que foi dito, atingindo o que não foi dito e, ainda, aquilo que poderia ter sido dito e que não foi. O discurso nas redes sociais vai além do texto escrito em telas, ofertando todas essas possibilidades aos/às usuários/as do ciberespaço de dizer o que pretendem, mas sem usar, necessariamente, a simbologia do alfabeto escrito. O discurso se realiza, assim, independentemente do uso de textos escritos. Outras maneiras de posicionamento estão a um clique do sujeito.

Para efeito de contextualização, sempre que nos referirmos a texto, no ambiente virtualizado do *Facebook*, estaremos nos direcionando a todas essas formas virtualizadas de manifestação de sentidos.

Para a Análise do Discurso (AD), o texto representa a unidade de análise, não importando sua extensão, nem se é um texto verbal ou não verbal. É no texto que o discurso se materializa e é nele que os aspectos histórico-sociais, ideológicos e subjetivos se solidificam (ORLANDI, 1999).

Quando nos referimos aos textos produzidos no *Facebook*, compreendemos como estes produzem sentido, levando-se em consideração, para isso, como se relacionam com os diferentes processos de significação e toda exterioridade. É necessário, também, entender que esse texto não se trata de um produto acabado, fechado. Ele sempre vai carregar consigo a memória discursiva e vai apontar para outro texto, para outro discurso. Tem-se, pois, um processo em curso, não sendo

12 Os *emojis* são figuras geradas para representar emoções humanas (PAIVA, 2016).

13 Meme são criações dos/as usuários/as da internet e que misturam uma situação que, em geral, obteve destaque na mídia, com diversas situações cotidianas e que ganham tom humorístico e irônico. Presente nas redes sociais, o meme é direcionado a situações de humor, mas, também, para críticas sociais, política e culturais. (GUERREIRO, SOARES, 2016).

14 *Hashtags* são palavras ou frases escritas precedidas do símbolo #. Servem para categorizar publicações, indexando conteúdos e tornando-os aptos para busca. As *hashtags* são de livre criação pelo/a usuário/a (PINTO, THEODORO, OLIVEIRA, 2016).

um conjunto de texto, mas sim uma prática que traduz a sociedade na história atual: a sociedade conectada.

Por esse viés, os textos inseridos nessa rede social são a ponte de acesso aos discursos produzidos pelos sujeitos no ciberespaço. Como unidade discursiva, esses textos trazem em si espaços significantes e, como todo objeto simbólico, são objetos de interpretação. Além de lê-los, trazemos à tona os sentidos que estão presentes e as relações de poder que ali se fazem latentes. Relações estas que são potencializadas devido à fácil e intensa relação com a cultura e a sociedade proporcionada pelo ciberespaço.

Consideramos, ainda, que o sujeito que produz o discurso, nessa rede social, o faz enquanto usuários/as do ciberespaço. Um mesmo sujeito, desse modo, pode desenvolver diferentes discursos a depender do lugar de fala que ocupa, mais precisamente o *Facebook*.

5 Considerações finais

Quando nos propomos a compreender o Facebook como espaço discursivo que se propaga através das conexões cibernéticas, identificamos que os sujeitos, enquanto usuári@s do Facebook, produzem seus discursos frente às diversas temáticas a partir de um lugar ideológico que, substancialmente, é forjado na convivência em sociedade. Dessa forma, impõem-se limites e fronteiras, sobretudo no que se refere à diversidade, materializados (esses discursos), também, por entre as redes da internet, nos dizeres que produzem exclusão aos/às diferentes.

Estes/as usuári@s, regidos/as pela memória discursiva, atrelados/as ao seu contexto histórico e ideológico, constroem, seja pelo texto verbal ou não verbal, significados de exclusão àqueles e àquelas que "ousam" quebrar o paradigma das "diferenças". Consideramos, pois, que os vestígios históricos e de produção dos discursos de exclusão dessas diferenças circulam, se mantêm e se propagam nas conexões cibernéticas, nas ditas comunidades virtuais.

Os discursos operam em redes que mantêm uma relação com os sujeitos e suas condições de produção e formatam uma maneira de pensar essas diferenças possibilitadas pelo mundo virtual de forma muito mais fortalecida pelo anonimato.

Para além dessa análise, se bem utilizadas, as redes sociais podem servir de ferramenta, inclusive, para dar voz aos/às mesmos/as usuári@s negligenciados/as e excluídos/as socialmente e promover o debate sobre uma diversidade de temas; assim, ajudam no compartilhamento de uma cultura de respeito às diferenças, sociais, raciais, religiosas, sexuais, por fim, de respeito à diversidade humana.

Referências

AMORIM, T. C. G. W. **Amizades das redes**: uma cartografia das amizades virtuais no Facebook. 2011. Dissertação (Mestrado Psicossociologia de Comunidades e Ecologia Social) - Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2011.

CANAL TECH. **Facebook revela relatório financeiro com números acima do esperado**. 2016. Disponível em: https://canaltech.com.br/redes-sociais/facebook-revela-relatorio-financeiro-com-numeros-acima-do-esperado-64061/. Acesso em: 18 jan. 2018.

CASTELLS, M. **A Sociedade em rede**. São Paulo: Paz e Terra, 2016.

GLOBO. **Facebook atinge os 2 bilhões de usuários**. 2017. Disponível em: http://g1.globo.com/tecnologia/noticia/facebook-atinge-os-2-bilhoes-de-usuarios.ghtml. Acesso em: 15 set. 2017.

GONDIM, P. G. C. **As construções discursivas de identidades no território do Facebook e as políticas educacionais de diversidade**. 2017. Tese (Doutorado em Educação) – Universidade Federal da Paraíba, João Pessoa, 2017.

GUERREIRO, A.; SOARES, N. M. M. Os memes vão além do humor: uma leitura multimodal para a construção de sentidos. **Texto Digital**, Florianópolis, v. 12, n. 2, p. 185-208, jul./dez. 2016. Disponível em: https://periodicos.ufsc.br/index.php/textodigital/article/view/1807-9288.2016v12n2p185. Acesso em: 15 set. 2017.

LEMOS A.; LÉVY, P. **O futuro da internet**. São Paulo: Paulus, 2010.

LÉVY, P. **Cibercultura**. São Paulo. Editora 34, 2010.

MARTINO, L. M. S. **Teoria das mídias digitais**: linguagens, ambientes e redes. Petrópolis: Vozes, 2015.

MAZZOCATO, S. B. **A reconfiguração do sujeito através de sua representação online**: as características e os processos no Facebook. 2014. Tese (Doutorado em Comunicação) – Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, 2014.

OLHAR DIGITAL. **Facebook já vale mais que Amazon e Oracle**. Disponível em: https://olhardigital.com.br/pro/noticia/facebook-ja-vale-mais-que-a-amazon-e-que-a-oracle/40336/. Acesso em: 18 jan. 2018.

ORLANDI E. P. **Análise de discurso**: princípios e procedimentos. Campinas: Pontes; 1999

PAIVA, V. L. M. O. A linguagem dos emojis. **Trabalhos em Linguística Aplicada**, Campinas, v. 55, n. 2, p. 379-399, maio/ago. 2016.

PENA, K. D. **Configurações do racismo nas redes sociais**. 2017. Dissertação (Mestrado Profissional em Gestão das Organizações Aprendentes) – Universidade Federal da Paraíba, João Pessoa, 2017.

PINTO, P.; THEODORO, I.; OLIVEIRA, J. Comportamento das hashtags durante grandes eventos. In: WORKSHOP SOBRE ASPECTOS DA INTERAÇÃO HUMANO-COMPUTADOR PARA A WEB SOCIAL, 7. 2016, São Paulo. **Anais** [...]. São Paulo, [s.n.] 2016.

TECNOBLOG. **Alternativa ao "curtir"**: Facebook lança botão de reações ao mundo inteiro. 2016. Disponível em: https://tecnoblog.net/192040/facebook-reacoes-botao-global/. Acesso em: 2 set. 2017.

TOMAÉL, M. I.; ALCARÁ. A. R.; DI CHIARA, I. G. Das redes sociais à inovação. **Ci. Inf.**, Brasília, v. 34, n. 2, p. 93-104, maio/ago. 2005.

Disponível em: http://www.scielo.br/pdf/ci/v34n2/28559.pdf/. Acesso em: 3 jan. 2017.

VICENTE, E. **Redes sociais**. Santa Cruz do Rio Pardo: Editora Viena, 2014.

Verbete, entidade e *query*: em busca de um construto do conhecimento e da significação no ciberespaço

Silvana Drumond Monteiro
Gian Carlo Decarli

1 Introdução

O capítulo, que ora se inicia, tem como objetivo demonstrar o construto do conhecimento e da significação que se estabelece no ciberespaço, a partir do percurso do **verbete**, na Wikipédia, **da entidade**, em banco de conhecimento como a Wikidata e a DBpedia, e da ***query*** (ou palavra-chave ou ainda de busca) no Google e Bing. Todo esse *continuum* semântico e semiótico estabelece possibilidades para estudo do fenômeno, em nível descritivo, para designar possíveis elaborações teóricas.

Tendo as palavras "verbete-entidade-*query*" definições próprias, o que se pretende é alcançar um nível de abstração que possa explicar, brevemente, a relação que se estabelece entre elas a partir do percurso da palavra na formação do conhecimento e da significação no ambiente ciberespacial.

Dessa forma, a palavra "constructo" atende os nossos objetivos, pois

> *Constructos* são abstrações de níveis mais elevados. Constituem uma edificação mais ideativa, abstrata, obtida a partir da agregação desde níveis de abstração mais simples até mais complexos. São verdadeiras 'montagens' e, portanto, decomponíveis. Representam como que uma edificação arquitetônica onde diversos materiais, compreendendo simples e compostos, são deliberadamente agregados com um objetivo definido. Os constructos visam a sumarizar fatos observáveis e a propor explicações. São a pedra de toque na interpretação de certos dados empíricos e na elaboração de teorias (ARY; JACOBS; RAZAVIEH, 1979, p. 26, grifos dos autores).

Para discorrer o tema, será necessário percorrer a cartografia da Web, isto é, suas fases que, não coincidentemente, são as mesmas das dimensões da linguagem: sintática, semântica e pragmática. São fenômenos do signo, portanto da significação e do conhecimento, embora outras denominações possam ser dadas, como Web das pessoas, Web dos dados, Web das coisas, Web local, Wel ubíqua e Web dos saberes, de acordo com Lévy (2014).

Assim, podemos esboçar alguns termos que despontam na literatura científica, que podem traduzir esse evento, como "ecossistema do conhecimento", denominação de Vanderdorpe (2015) para atribuir à Wikipédia como parte dele, ou então "*Web* Pragmática" (DECARLI, 2017; DECARLI; MONTEIRO, 2017), para designar a convergência das *Webs* e a atribuição do sentido no ciberespaço, ou, até mesmo, em certa medida, "esfera semântica" (LÈVY, 2014). O fato é que os fenômenos só se tornam cientificamente inteligíveis quando uma teoria os filtra e demonstra as suas relações.

Mesmo que, conforme Freitas (1994), os constructos sejam substituídos por sínteses mais eficazes ou por explicações mais precisas para os fenômenos que representam, é necessário que a Ciência da Informação e, portanto, os profissionais da informação desafiem a nova realidade que se apresenta nas plataformas digitais que envolvam o conhecimento.

2 Um possível trajeto para assimilação do sentido na *web*

A *World Wide Web* (WWW ou *Web*), interface interoperável de metarrepresentação no ciberespaço, não é estática. Seu conteúdo, funcionamento e refinamento tecnológico têm de ser constante para atender a distintos contextos de uso. A *Web* indexa e organiza o conhecimento mundial. Em uma perspectiva prática, segundo Bax (2012, p. 8), a *web* é "[...] um espaço virtual construído sobre a estrutura física da Internet." Assim, em sua hipertextualidade, disponibiliza mecanismos que possibilitam filtrar e selecionar a informação disponível (POMBO, 2006).

A interoperabilidade é um conceito que a envolve desde sua gênese e transpassa todas as readequações tecnológico-estrutural realizadas ao longo de sua existência. Essas transformações podem ser delineadas com base em uma perspectiva linguística, especificamente, associada à tríade das disciplinas de

sintaxe, semântica e pragmática. Na Ciência da Informação, Vechiato e Vidotti (2014) se baseiam no conceito de dobra semiótica apresentado por Monteiro (2012) e também nas dimensões da linguagem explanadas por Jorente (2012) para designá-las como: *Web* Sintática (*Web* 1.0 e *Web* 2.0 ou Social); *Web* Semântica (3.0 ou de Dados); *Web* Pragmática (4.0 ou Social Semântica).

Acontecimentos como a ausência de mecanismos de busca sofisticados, o surgimento dos primeiros websites e a adoção de *e-mails*, *chats* e fóruns para a troca de informações caracterizam a fase designada como *Web* 1.0. Na *Web* 2.0 ou colaborativa, sua sucessora, há o aparecimento das redes sociais, *Wikis* e o crescimento de mecanismos de busca como o Google (SANTAELLA, 2013). A "*Web* Sintática" abrange desde a estrutura estática dos *websites* até as práticas colaborativas de comunidades *on-line* (VECHIATO; VIDOTTI, 2014).

Os *weblogs* e as *wikis* são duas importantes tecnologias desenvolvidas durante o período da *Web* Social e contribuíram significativamente para a expansão da cooperação. Nessa esteira, em 2001, a Wikipédia foi originada e seu funcionamento representa o propósito colaborativo da *Web* Social (AYERS; MATTHEWS; YATES, 2008). Na plataforma, sujeitos corroboram para criação, manutenção e atualização das informações veiculadas, tornando-a uma grande base internacional do conhecimento.

Pombo (2006) argumenta que nas enciclopédias eletrônicas (*on-line* ou CD-ROM), neste exemplo a Wikipédia, a leitura foi substituída pela navegação, a manipulação entre volumes pelos "cliques" e a linearidade pela possibilidade de percorrer diversos caminhos livremente. Ainda, segundo a autora, esses novos formatos de enciclopédias parecem disponibilizar o conhecimento necessário para a resolução de quase todas as perguntas possíveis. Assim, como em uma enciclopédia tradicional, a Wikipédia busca um nível de detalhamento maior do contexto que envolve seus verbetes e não apenas seu significado.

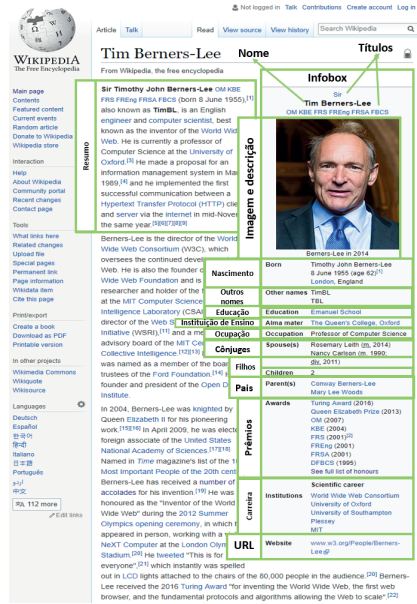
Para Croft e Cruse (2004), o significado, sob uma perspectiva humana de cognição, ocorre a partir de uma breve inferência descritiva de um conceito, complementado pelo conhecimento de mundo do sujeito. Koch e Elias (2008) corroboram ao expressar que os conhecimentos de mundo e enciclopédico são equivalentes, pois remetem a conhecimentos gerais, obtidos a partir de vivências pessoais como subsídios a produção de sentido. Nessa perspectiva, o

conhecimento enciclopédico está intimamente relacionado a fatores empíricos e ao saber contextual para além de descrições genéricas.

O conhecimento enciclopédico existente na Wikipédia é constituído, essencialmente, da colaboração entre seus membros, seja para a escrita de um conteúdo, formatação de seções, criação de *bots* (robôs) de monitoramento dos textos e até a criação de grupos com a indicação de responsáveis para o controle e validação das informações veiculadas nos artigos.

Na Figura 1, há um exemplo de verbete da Wikipédia, mais precisamente, o correspondente ao termo "Tim Berners-Lee", escolhido em razão de sua representatividade acerca do tema *"Web"*. O *infobox* vinculado ao artigo sobre "Tim Berners-Lee" apresenta sucintas informações fatuais a respeito de acontecimentos ocorridos ao longo de sua vida. Ao todo, o *infobox* do verbete de "Tim Berners-Lee" é composto por onze campos identificados como: *"Born"*, *"Other names"*, *"Education"*, *"Alma mater"*, *"Occupation"*, *"Spouse(s), "Children"*, *"Parent(s), "Awards"*, *"Institutions"* e *"Website"*.

Figura 1 – Verbete do termo "Tim Berners-Lee" na Wikipédia



Fonte: Adaptado de Wikipédia (2018).

A maioria dos campos que compõe o *infobox* de "Tim Berners-Lee" possui *hyperlinks* que redirecionam o acesso ao conteúdo de um verbete na íntegra. Apenas os campos *"Other names"*, *"Occupation"*, *"Spouse(s)"* e *"Children"* não possuem verbetes criados na Wikipédia que permitam um redirecionamento interno.

Comumente, um artigo da Wikipédia é composto pelas guias *"Article/Read"*, *"Talk"*, *"View Source"* e *"View History"*. Nesta última, o aspecto colaborativo da Wikipédia é evidenciado, pois, ao acessá-la, é possível visualizar a relação de usuários, alterações e datas que compõem a construção do verbete ao longo do tempo.

A Wikipédia foi brevemente abordada para que fosse possível exemplificar o funcionamento colaborativo de uma tecnologia originada na *Web* Social. Entretanto, esse não foi o único motivo de sua aparição, uma vez que a Wikipédia é apenas um dos agentes responsáveis pelo ciclo de transformação do conhecimento na *Web*.

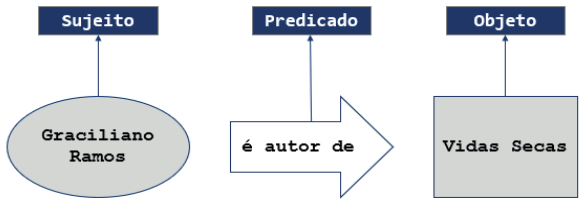
De volta às dimensões da *Web*, a *Web* Semântica ou 3.0 é uma extensão da anterior, na qual a informação tem seu significado bem definido para que pessoas e agentes computacionais possam trabalhar em cooperação (BERNERS-LEE; HENDLER; LASSILA, 2001). A relação humano-máquina, sugerida pelos autores, está além da mecanização e automatização de funções e tarefas, e consiste na capacidade do computador em compreender os devires da comunicação humana.

Para tanto, sujeitos e agentes computacionais apenas compartilharão de um entendimento acerca de um objeto se houver a formalização do sentido. Na *Web* Semântica, um conglomerado de tecnologias é responsável por essa tarefa de padronização do significado. Entretanto, neste capítulo, um dos enfoques está no funcionamento de uma dessas tecnologias, as *Knowledge Bases* que, segundo Uyar e Aliyu (2015), são responsáveis por estruturar, em seu âmago, as informações dispersas na *Web* através de conexões semânticas entre entidades.

As *Knowledge Bases* utilizam, principalmente, o *Resource Description Framework* (RDF) para o enriquecimento semântico de informações. Seu emprego permite descrever coisas do mundo, tal como pessoas, locais e conceitos complexos, como também de que maneira estes se relacionam. A origem do RDF, segundo Bax (2012), está ligada aos estudos de redes semânticas realizados pela área de Inteligência Artificial (IA). Porém, difere-se ao passo que é uma linguagem de representação do conhecimento utilizada para que agentes computacionais

tenham capacidade de processar suas informações por meio de uma semântica formalizada, conforme exposto na Figura 2.

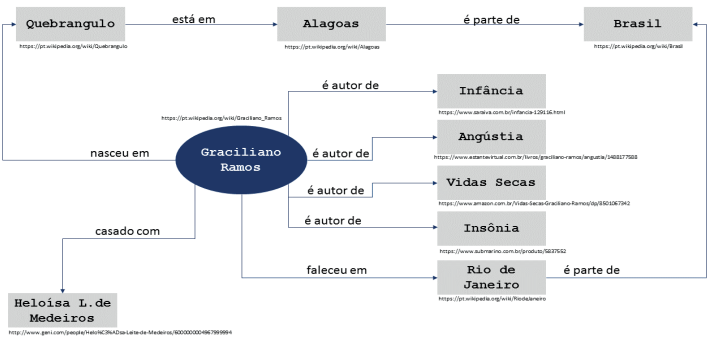
Figura 2 – Modelo de declaração RDF em triplas: sujeito, predicado e objeto



Fonte: Baseado em Bax (2012).

Desse modo, a descrição de um recurso é feita por meio das triplas RDF, as quais têm a função de espelhar a estrutura básica de uma frase e são formadas por sujeito, predicado e objeto (HEATH; BIZER, 2011). As *Knowledge Bases* extraem informações estruturadas para que possam ser criadas representações a partir delas, mediante a aplicação de triplas RDF (sujeito, predicado e objeto), apresentadas no formato de declarações. O cruzamento dessas informações permite o relacionamento entre entidades, que resulta em respostas mais concisas para quando se necessita saber qual a cidade de nascimento de uma celebridade ou em quais filmes já atuou (BIEGA; KUZEY; SUCHANEK, 2013).

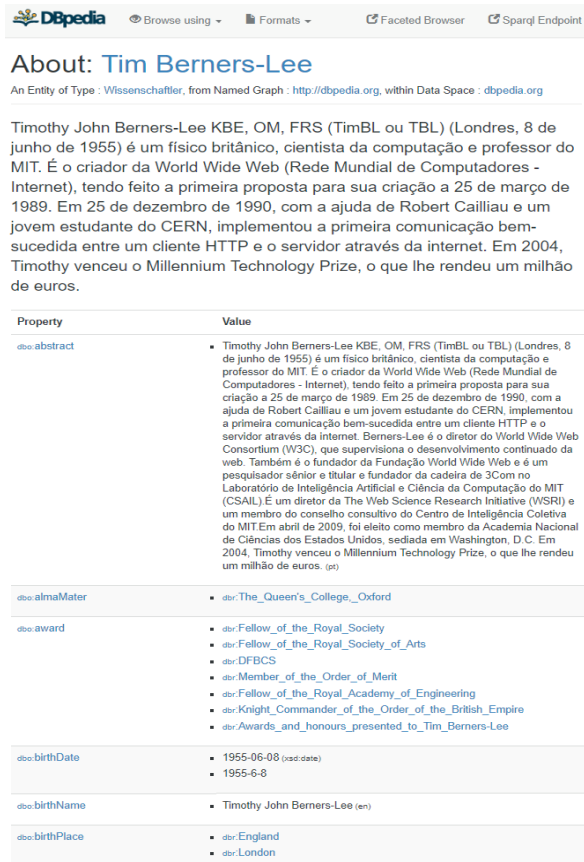
Figura 3 – Modelo de grafo RDF: fusão de triplas



Fonte: Autoria própria (2018).

Na Figura 3, há um exemplo de como ocorre, conceitualmente, a união das triplas RDF para a formação de *datasets*. Segundo Bax (2012), as *Uniform Resource Identifier* (URI) vinculadas às triplas RDF são compartilhadas com outros conjuntos de dados para a formação de grafos que, sucessivamente, podem se unir a outros grafos, deixando-os cada vez maiores. Na prática, ao acessar a página correspondente ao termo "Tim Berners-Lee", no *website* da DBpedia, é possível observar o modo como as informações, antes veiculadas na Wikipédia, transformaram-se em declarações, conforme exposto na Figura 4.

Figura 4 – O termo "Tim Berners-Lee" na DBpedia



Fonte: DBpedia (2018).

Os valores que preenchem as propriedades da DBpedia são equivalentes aos encontrados no artigo da Wikipédia sobre o termo "Tim Berners-Lee", especialmente, os apresentados no *infobox* da Wikipédia. Por exemplo, a propriedade "dbo: almaMater" exibe a mesma informação contida no campo "Alma Mater" do *infobox*. Pode-se inferir que as informações advindas da Wikipédia são semanticamente enriquecidas por meio de declarações (sujeito, propriedade, objeto) na DBpedia.

Nas triplas RDF, segundo Bax (2012), sujeito e objeto são entes e a propriedade corresponde à relação existente entre esses entes, sendo que todos (sujeito, propriedade e valor) são representados por uma URI. Assim, as declarações promovem a construção do significado formal para agentes computacionais na *Web*, mediante os relacionamentos estabelecidos entre as entidades com o uso de URI.

No Wikidata os verbetes são convertidos em itens, propriedades e/ou valores. As entidades são compreendidas como *"Itens"*, também compostos por propriedades e valores, sendo que a propriedade é uma categoria de dados e o valor é a própria descrição do *"Item"* (WIKIDATA, 2018a). Diferente do modelo de *website* adotado pelo DBpedia, o Wikidata possui muitas semelhanças de interface e funcionamento com a Wikipédia, em virtude de compartilharem da tecnologia e ideologia *"wiki"* no desenvolvimento e manutenção dessas plataformas por sua mantenedora, a Wikimedia Foundation.

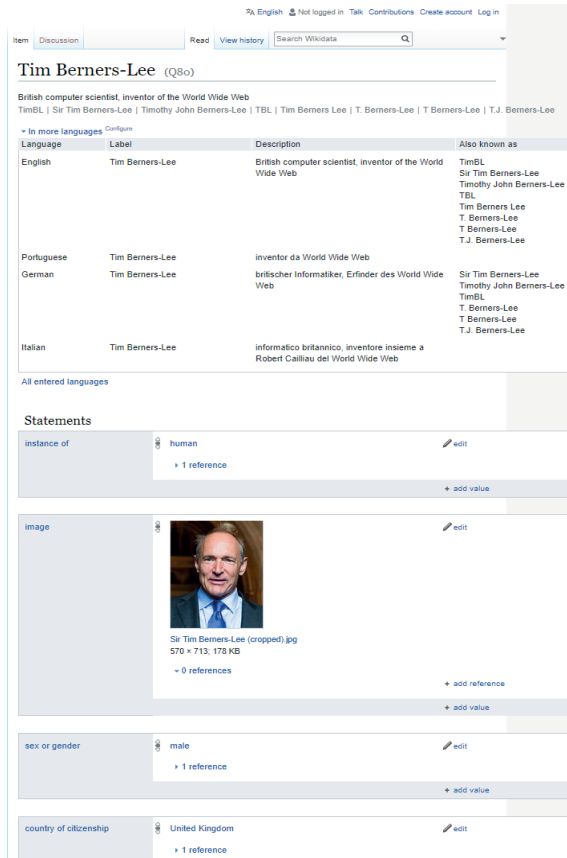
Por meio de procedimentos semelhantes aos da DBpedia, o Wikidata extrai as informações da Wikipédia e também fornece dados para a complementação de seus verbetes, isto é, há uma relação inversa, pois ao mesmo tempo que extrai também fornece informações, conforme o interesse da comunidade (WIKIDATA, 2018a).

A página do *"Item"* também possui guias semelhantes às usadas na Wikipédia, sendo formada pela sequência: *"Item/Read"*, *"Discussion"*, *"View source"* e *"View history"*. Na primeira é possível visualizar e editar as declarações da entidade; em *"Discussion"* podem ser abertos tópicos sobre melhorias/correções de conteúdo; na terceira a fonte pode ser visualizada e editada; e em *"View history"*

é exibido o histórico de revisões e os usuários responsáveis, de maneira idêntica ao funcionamento da Wikipédia.

Os *"Itens"* são identificados de duas formas: a primeira, chamada de *"Label"*, é destinada a visualização por humanos e remete ao próprio nome "Tim Berners-Lee"; a segunda é denominada de *"Unique identifier"* e corresponde a uma codificação do texto que identifica o *"Item"* para que seja facilmente processado por máquinas. No caso de "Tim Berners-Lee" é usado o "Q80", conforme exposto na Figura 5.

Figura 5 – O termo "Tim Berners-Lee" no Wikidata



Fonte: Wikidata (2018b).

As informações existentes no *"Item"* de "Tim Berners-Lee" advêm de diferentes fontes, como *hyperlinks* para *websites* externos, além de Wikipédia, Wikisource, Wikibooks, Wikiquote, entre outros projetos da Wikimedia Foundation. A relação direta com a Wikipédia é evidente, porém não há como precisamente indicar seu fluxo de extração de dados. Na segunda seção do *"Item"* há os agrupamentos de *"Statements"*, nos quais estão alocadas as *"Property"* e o *"Value"*, além de *"Qualifiers"* para declarações adicionais sobre os próprios valores.

As tecnologias da *Web* Semântica estão presentes na maioria dos processos que envolvem a interoperabilidade desses dados no âmago dessas *Knowledge Bases*, seja para a padronização, extração e/ou publicação de informações. Apresentar, ainda que brevemente, a forma como as informações são organizadas no DBpedia e Wikidata torna possível compreender como ocorre o processo de formalização do significado no contexto da *Web*.

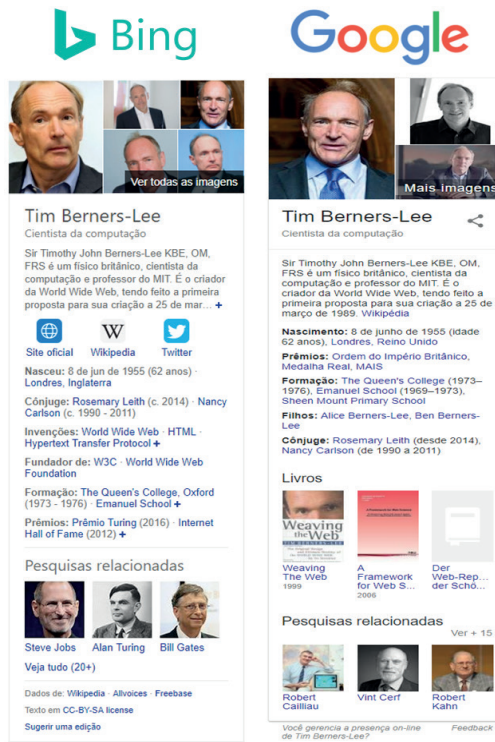
Em continuidade, a pragmática aplicada à *Web* tem o mesmo sentido compartilhado pelas áreas de Linguística e Semiótica: o de buscar compreender como o contexto influencia na interpretação do significado. Dessa maneira, em uma comunidade de colaboração, os significados individuais compõem uma estrutura cognitiva que está em sintonia com outras por meio de um processo comunicativo. Isso permite que a máquina possa lidar com questões de ambiguidade, semântica, relevância e sobrecarga de informações (PASCHKE *et al*., 2007).

Na *Web* Pragmática, recursos semânticos são colocados em contextos pragmáticos, os quais devem compor-se da junção entre conceitos comuns (senso comum) e individuais para que faça sentido sua adoção na comunidade (MOOR, 2005). Santaella (2012, p. 41) argumenta que na *Web* Pragmática

> [...] as páginas não serão acessadas somente por seu significado, mas também por sua utilização, como, por exemplo, pelas características dos grupos sociais e políticos de que o usuário participa, ou pelo perfil de pesquisa acadêmica a que pertence etc.

A *Web* Pragmática corresponde à *Web* Semântica em seu plano ideal, como já afirmado, pois utiliza a mesma estrutura semântica com o propósito de disponibilizar o significado pautado no consenso para sujeitos em múltiplos contextos. Nesse sentido, os atuais mecanismos de busca estão cada vez mais engajados sob a concepção de uma *Web* Pragmática. Além disso, o *Knowledge Graph*, tecnologia presente na página de exibição de resultados dos principais mecanismos de busca, auxilia na estruturação e apresentação de determinadas informações fatuais, conforme exposto na Figura 6. O *Knowledge Graph* é uma tecnologia recente e "[...] representa um grande avanço no tratamento e na busca por informações sobre entes do mundo real, pessoas, eventos, coisas etc." (BAX, 2012, p. 6).

Figura 6 – *Knowledge Graph* sobre "Tim Berners-Lee" em Bing e Google



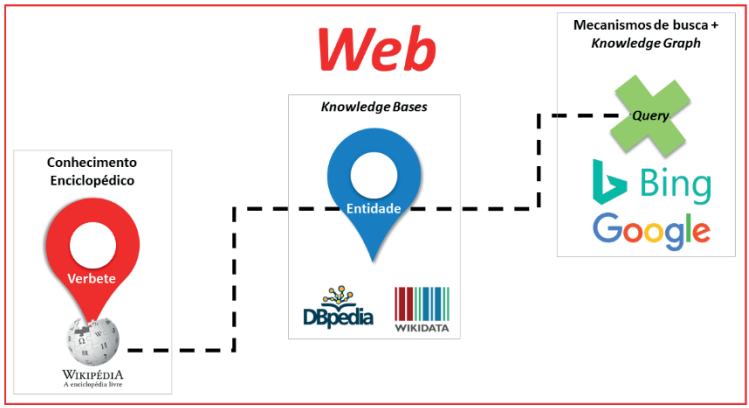
Fonte: Adaptado de Bing (2018) e Google (2018).

Na Figura 6, o trajeto percorrido pela informação é evidenciado, visto que os *Knowledge Graphs* dos mecanismos de busca Bing e Google apresentam uma disposição de dados similar àquelas contidas no *infobox* da Wikipédia e também nas entidades existentes na DBpedia e Wikidata para o termo "Tim Berners-Lee".

Onde há os campos que descrevem informações fatuais sobre "Tim Berners-Lee", os valores são *hyperlinks* que redirecionam para a página de seus relacionamentos por meio de requisições prontas (*query's*) para cada item, isto é, os valores que constam na entidade de "Tim Berners-Lee" são compostos por outras páginas de exibição de resultados.

O verbete convertido em entidade nas *Knowledge Bases* mais tarde se tornou resultado nos mecanismos de busca. Desse modo, as páginas de exibição de resultados de Bing e Google foram exploradas no intuito de verificar se este conhecimento originado na Wikipédia e tratado por Wikidata e DBpedia era utilizado na apresentação dos resultados, conforme exposto na Figura 7.

Figura 7 – Um caminho provável para a formalização do significado na *Web*



Fonte: Autoria própria (2018).

Equivalente ao percurso traçado (de um ponto A para um ponto B) em um aplicativo de navegação ou aparelho GPS, a *Web* também possui diferentes caminhos a serem percorridos para a construção do mútuo significado para sujeitos e agentes computacionais em seu âmago. Ressalta-se que a rota (verbete,

entidade e *query*), explorada nesta investigação, não é o único trajeto, mas um caminho possível para o estabelecimento do sentido nesse contexto.

3 Considerações finais

A forma como essas informações são dispostas e exibidas é apenas o produto de um longo processo de estruturação de dados na *Web*. O percurso é iniciado na Wikipédia por meio de um verbete redigido colaborativamente que, posteriormente, será submetido a um enriquecimento semântico em *Knowledge Bases* como Wikidata e DBpedia. Por fim, mecanismos de busca como Bing e Google as utilizam em suas páginas de exibição de resultados.

Não há como negar a importância da Wikipédia na ecologia do conhecimento no ciberespaço. Vanderdorpe (2015) alerta que as enciclopédias, através da história, têm sido não somente repositórios do conhecimento, mas também lugares que oferecem uma reflexão sobre a natureza do conhecimento e sua organização.

Em virtude de sua representatividade acerca do tema, o termo "Tim Berners-Lee" foi elencado no intento de vislumbrar a trajetória do conhecimento enciclopédico na *Web*. Na primeira fase, esse termo consultado na Wikipédia, sob a condição de verbete; na segunda etapa, submetido à verificação na DBpedia e Wikidata enquanto entidade. No terceiro estágio, o termo foi inserido nas caixas de pesquisa dos seguintes mecanismos de busca: Bing e Google.

Com base nessa conjuntura tecnológica, este capítulo demonstrou como tecnologias, aparentemente distintas, confluem sinergicamente para a instauração da *Web* Pragmática. Assim, seu enfoque esteve sobre as tecnologias desenvolvidas na *Web* Social, potencializadas na *Web* Semântica e que, consequentemente, interconectaram-se na *Web* Pragmática, envolvidas com a organização, disseminação e recuperação de informações nesse domínio.

O conceito "ecologia do conhecimento" tem sido usado para descrever situações em que múltiplos atores têm relacionamentos dinâmicos. É certo que o tema é multifacetado e implica em múltiplas ressonâncias na Ciência da Informação, no entanto, e por isso mesmo, faz-se *mister* os profissionais dessa

área pesquisá-lo, e nesse sentido, e de acordo com Szonieck (2018), a Sociedade Francesa de Ciências da Informação e da Comunicação (SFIC) o tem utilizado para o campo das humanidades digitais para abordar os ecossistemas complexos.

Referências

ARY, D.; JACOBS, L.C.; RAZAVIEH, A. **Introduction to research in education**. New York: Holt, Heinehart & Wilston, 1979.

AYERS, P.; MATTHEWS, C.; YATES, B. **How Wikipedia works**. San Francisco: No Starch Press, 2008.

BAX, M. P. A evolução da web rumo à web semântica. **PRISMA.COM**, Porto, n. 19, p. 1-27, 2012.

BERNERS-LEE, T.; HENDLER, J.; LASSILA, O. The semantic web: a new form of Web content that is meaningful to computers will unleash a revolution of new possibilities. **Scientific American**, maio 2001. Disponível em: http://cogweb.ucla.edu/crp/Media/2001-05-05_SemanticWeb.html. Acesso em: 28 maio 2018.

BIEGA, J.; KUZEY, E.; SUCHANEK, F. Inside YAGO2s: a transparent information extraction architecture. *In*: THE DEMO FALL, 1., 2013, Silicon Valley. **Proceedings** [...]. Silicon Valley: Demo Fall, 2013. Disponível em: https://suchanek.name/work/publications/www2013demo.pdf. Acesso em: 28 maio 2018.

BING. Disponível em: https://www.bing.com/. Acesso em: 28 maio 2018.

CROFT, W.; CRUSE, D. A. **Cognitive linguistics**. New York: Cambridge University Press, 2004.

DBPEDIA. **Página sobre Tim Berners-Lee**. 2018. Disponível em: http://dbpedia.org/page/Tim_Berners-Lee. Acesso em: 28 maio 2018.

DECARLI, G. C. **O conhecimento enciclopédico colaborativo e suas influências no desenvolvimento da web pragmática**. 2017. Dissertação (Mestrado em Ciência da Informação) – Universidade Estadual de Londrina, Londrina, 2017. Disponível em: http://www.bibliotecadigital.uel.br/document/?code=vtls000214108. Acesso em: 28 maio 2018.

DECARLI, G. C.; MONTEIRO, S. D. Knowledge graph: o elo entre o conhecimento enciclopédico colaborativo e a web pragmática. *In*: ENCONTRO NACIONAL DE PESQUISA E PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO, 18., 2017, Marília, **Anais** [...]. Marília: Unesp, 2017. Disponível em: http://enancib.marilia.unesp.br/index.php/xviiienancib/ENANCIB/paper/view/122. Acesso em: 28 maio 2018.

FREITAS, E. L. Alguns aspectos da linguagem científica. **Sitientibus**, n. 12, p. 101-112, 1994. Disponível em: http://www2.uefs.br/sitientibus/pdf/12/alguns_aspectos_da_linguagem_cientifica.pdf. Acesso em: 28 maio 2018.

GOOGLE. Disponível em: https://www.google.com.br/. Acesso em: 28 maio. 2018.

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

JORENTE, M. J. V. **Ciência da informação**: mídias e convergência de linguagens na web. São Paulo: Cultura Acadêmica, 2012.

KOCH, I.; ELIAS, V. **Ler e compreender**: os sentidos do texto. São Paulo: Contexto, 2008.

LÉVY, P. **A esfera semântica**: computação, cognição, economia da informação. São Paulo: Annablume, 2014.

MONTEIRO, S. D. A dobra semiótica e os agenciamentos maquínicos: por uma ontologia das tecnologias de informação e comunicação.

In: CERVANTES, B. M. N. (org.). **Horizontes da organização da informação e do conhecimento**. Londrina: EDUEL, 2012. p. 63-96.

MOOR, A. de. Patterns of the pragmatic web. *In*: INTERNATIONAL CONFERENCE ON CONCEPTUAL STRUCTURES, 13., 2005, Kassel. **Proceedings** [...]. Berlin: Kassel, 2005. Disponível em: https://link.springer.com/chapter/10.1007/11524564_1. Acesso em: 28 maio 2018.

PASCHKE, A. *et al*. Rule responder: ruleML-based agents for distributed collaboration on the pragmatic web. *In*: PROCEEDINGS INTERNATIONAL CONFERENCE ON THE PRAGMATIC WEB, 2., 2007, Tilburg. **Proceedings** [...]. Tilburg: ACM, 2007. Disponível em: http://ruleml.org/papers/RuleResponder.pdf. Acesso em: 28 maio 2018.

POMBO, O. O hipertexto como limite da ideia de enciclopédia. *In*: POMBO, O.; GUERREIRO, A.; ALEXANDRE, A. F. **Enciclopédia e hipertexto**. Lisboa: Edições Duarte Reis, 2006.

SANTAELLA, L. A tecnocultura atual e suas tendências futuras. **Signo y Pensamiento**, Bogotá, v. 30, n. 60, p. 30-43, jan./jun. 2012. Disponível em: http://revistas.javeriana.edu.co/index.php/signoypensamiento/article/download/2408/1692. Acesso em: 28 maio 2018.

SANTAELLA, L. Desafios da ubiquidade para a educação. **Revista Ensino Superior**, Campinas, n. 9, v. 2, p. 19-28, abr./jun. 2013. Disponível em: https://www.revistaensinosuperior.gr.unicamp.br/edicoes/edicoes/ed09_abril2013/NMES_1.pdf. Acesso em: 28 maio 2018.

SZONIECKY, S. **Ecosystems knowledge**: modeling and analysis method for information and communication. London: Wiley, 2018.

UYAR, A.; ALIYU, F. M. Evaluating search features of Google Knowledge Graph and Bing Satori: entity types, list searches and query interfaces. **Online Information Review**, n. 1, v. 39, p. 1-26, jan. 2015. Disponível

em: https://www.emeraldinsight.com/doi/abs/10.1108/OIR-10-2014-0257. Acesso em: 28 maio 2018.

VANDENDORPE, C. Wikipedia and the ecosystem of Knowledge. **Scholarly and Research Communication**, v. 6, n. 3, p.1-12, 2015. Disponível em: http://src-online.ca/index.php/src/article/download/201/423. Acesso em: 28 maio 2018.

VECHIATO, F. L.; VIDOTTI, S. A. B. G. Encontrabilidade da informação: atributos e recomendações para ambientes informacionais digitais. **Informação e Tecnologia**, João Pessoa, v. 1, n. 2, p. 42-58, jul./dez. 2014. Disponível em: http://periodicos.ufpb.br/ojs/index.php/itec/article/view/22099. Acesso em: 28 maio 2018.

WIKIDATA. **Página inicial**. Disponível em: https://www.wikidata.org/. Acesso em: 28 maio 2018b.

WIKIDATA. **Página sobre Tim Berners-Lee**. Disponível em: https://www.wikidata.org/wiki/Q80. Acesso em: 28 maio 2018a.

WIKIPÉDIA. **Página sobre Tim Berners-Lee**. Disponível em: https://en.wikipedia.org/wiki/Tim_Berners-Lee. Acesso em: 28 maio 2018.

Universidades que participam desta obra

Centro Universitário Eurípides de Marília (UNIVEM) - Brasil
Universidad de Murcia (UM) – Espanha
Universidad del Cauca (Unicauca) – Colombia
Universidad Pompeu Fabra (UPF) – Espanha
Universidade de São Paulo (USP) - Brasil
Universidade do Minho (UMinho) – Portugal
Universidade Estadual da Paraíba (UEPB) - Brasil
Universidade Estadual de Londrina (UEL) - Brasil
Universidade Estadual Paulista "Júlio de Mesquita Filho" (UNESP) - Brasil
Universidade Federal da Paraíba (UFPB) - Brasil
Universidade Federal de Pernambuco (UFPE) - Brasil
Universidade Federal de Santa Catarina (UFSC) - Brasil
Universidade Federal do Rio Grande do Norte (UFRN) - Brasil

Autorxs

Caio Saraiva Coneglian (UNESP, Brasil)
Doutorando em Ciência da Informação na Universidade Estadual Paulista (PPGCI/UNESP). Mestre em Ciência da Informação pela Universidade Estadual Paulista (UNESP), 2017. Graduado em Ciência da Computação pelo Centro Universitário Eurípedes de Marília (UNIVEM), 2014. Docente do Centro Universitário Eurípedes de Marília (UNIVEM).

Carlos Alberto Baptista Sousa Pinto (UMinho, Portugal)
Doutor em Ciências e Tecnologias da Comunicação pela Universidade de Aveiro (UA). Mestre em Informática de Gestão pela Universidade do Minho (UMinho). Licenciado em Engenharia de Sistemas e Informática pela Universidade do Minho (UMinho). Professor Auxiliar do Departamento de Sistema de Informação na Escola na Engenharia da Universidade do Minho - Guimarães (DSI/UMinho).

Daniel Thiago Freire da Silva (UFPB, Brasil)
Mestre em Gestão das Organizações Aprendentes pela Universidade Federal da Paraíba (UFPB), 2018. Especialista em Secretariado Executivo pela AVM Faculdade Integrada, 2014. Graduado em Secretariado pela Universidade Federal de Pernambuco (UFPE), 2011. Secretário Executivo da Universidade Federal de Pernambuco (UFPE).

Edgar Bisset Alvarez (UFSC, Brasil)
Doutor em Ciência da Informação pela Universidade Estadual Paulista (UNESP), 2017. Mestre em Biblioteconomia e Ciência da Informação pela Universidade da Havana (2012). Graduado em Informação Técnico-Científica e Biblioteconomia pela Universidade de Havana (2007). Professor do Departamento de Ciência da Informação da Universidade Federal de Santa Catarina.

Fernando Luiz Vechiato (UFRN)
Doutor em Ciência da Informação pela Universidade Estadual Paulista (UNESP), 2013. Mestre em Ciência da Informação pela Universidade Estadual Paulista (UNESP), 2010. Bacharel em Biblioteconomia pela Universidade Estadual Paulista (UNESP), 2007. Professor Adjunto do Departamento de Ciência da Informação da Universidade Federal do Rio Grande do Norte (DCI/UFRN) e Professor Colaborador do Programa de Pós-Graduação em Ciência da Informação da Universidade Estadual Paulista (PPGCI/UNESP).

Gian Carlo Decarli (UEL, Brasil)
Doutorando no Programa de Pós-Graduação em Ciência da Informação da Universidade Estadual Paulista (PPGCI/UNESP). Mestre em Ciência da Informação pelo Programa de Pós-Graduação em Ciência da Informação da Universidade Estadual de Londrina (UEL), 2017. Graduado em Biblioteconomia pela Universidade Estadual de Londrina (UEL), 2012.

Henry Poncio Cruz de Oliveira (UFPB, Brasil)
Doutor em Ciência da Informação pela Universidade Estadual Paulista (UNESP), 2014. Mestre em Ciência da Informação pela Universidade Federal da Paraíba (UFPB), 2010. Graduado em Física pela Universidade Federal da Paraíba (UFPB), 2007. Técnico em Eletrônica pela Escola Técnica Federal da Paraíba (atual IFPB), 1996. Professor Adjunto do Departamento de Ciência da Informação da Universidade Federal da Paraíba (DCI/UFPB). Docente do Programa de Pós-Graduação em Ciência da Informação da Universidade Federal da Paraíba (PPGCI/UFPB).

Izabel França de Lima (UFPB, Brasil)
Doutora em Ciência da Informação pela Universidade Federal de Minas Gerais (UFMG), 2012. Mestre em Educação pela Universidade Federal da Paraíba (UFPB), 2007. Especialista em Gestão de Unidades de Informação pela Universidade Federal da Paraíba (UFPB), 2006. Graduada em Biblioteconomia (1989) e em Administração (1999) pela Universidade Federal da Paraíba (UFPB). Professora

Adjunta do Departamento de Ciência da Informação da Universidade Federal da Paraíba (DCI/UFPB). Docente no Programa de Pós-Graduação em Ciência da Informação da Universidade Federal da Paraíba (PPGCI/UFPB).

José Eduardo Santarem Segundo (USP, Brasil)
Doutor em Ciência da Informação pela Universidade Estadual Paulista (UNESP), 2010. Mestre em Ciência da Informação pela Universidade Estadual Paulista (UNESP), 2006. Especialista em Tecnologia e Desenvolvimento de Sistemas para Internet pela Universidade de Marília (UNIMAR), 2002. Graduado em Tecnologia em Processamento de Dados pela Universidade de Marília (UNIMAR), 2000. Professor do Departamento de Educação, Informação e Comunicação da Universidade de São Paulo (DEIC/USP). Docente no Programa de Pós-Graduação em Ciência da Informação da Universidade Estadual Paulista "Júlio de Mesquita Filho" (PPGCI/UNESP).

Josevânia da Silva (UEPB, Brasil)
Doutora em Psicologia (Psicologia Social) pela Universidade Federal da Paraíba (UFPB), 2011. Mestre em Psicologia Social pela Universidade Federal da Paraíba (UFPB), 2009. Licenciada em Psicologia pela em pela Universidade Federal da Paraíba (UFPB), 2008. Psicóloga pela Universidade Federal da Paraíba (UFPB). 2008. Docente Adjunta do Departamento de Psicologia da Universidade Estadual da Paraíba (DP/UEPB). Professora no Programa de Pós-graduação em Psicologia da Saúde (Mestrado) da Universidade Estadual da Paraíba (PPGPS/UEPB).

Juan Antonio Pastor-Sánchez (UM, Espanha)
Doctor en documentación por la Universidad de Murcia. Profesor del Área de Biblioteconomía y Documentación de la Facultad de Comunicación y Documentación de la Universidad de Murcia.

Leonardo Castro Botega (UNIVEM, Brasil)
Pós-doutorando na Universidade de São Paulo (USP). Doutor em Ciência da Computação pela Universidade Federal de São Carlos (UFSCar), 2016. Mestre em Ciência da Computação pela Universidade Federal de São Carlos (UFSCar), 2008. Bacharel em Ciência da Computação pelo Centro Universitário Eurípides de Marília (UNIVEM), 2005. Docente do Centro Universitário Eurípides de Marília (UNIVEM). Docente no Programa de Pós-Graduação em Ciência da Informação da Universidade Estadual Paulista (PPGCI/UNESP).

Maria da Luz Olegário (UFPB, Brasil)
Doutora em Educação pela Universidade Federal da Paraíba (2010). Mestre em Língua Portuguesa pela Universidade Federal da Paraíba (2001). Especialista em Leitura e Produção de Textos pela Universidade Federal da Paraíba (1998). Graduada em Letras pela Universidade Federal da Paraíba (1997). Professora da Universidade Federal da Paraíba. Docente do Mestrado Profissional em Letras da Universidade Federal de Campina Grande (MPL-UFCG).

Mayane Paulino de Brito e Silva (UFPB, Brasil)
Doutoranda em Ciência da Informação na Universidade Federal da Paraíba (PPGCI/UFPB). Mestre em Ciência da Informação pela Universidade Federal da Paraíba (2017). Bacharela em Biblioteconomia pela Universidade Federal do Rio Grande do Norte (2016). Técnica em Informática pelo Instituto Federal do Rio Grande do Norte (2011).

Plácida Leopoldina Ventura Amorim da Costa Santos (UNESP, Brasil)
Livre-docente em Catalogação pela Universidade Estadual Paulista (UNESP), 2010. Doutora em Letras - Semiótica e Linguística Geral pela Universidade de São Paulo (USP), 1994. Mestre em Ciência da Informação pela Pontifícia Universidade Católica de Campinas (PUCCamp), 1983. Bacharel em Biblioteconomia pela Universidade Estadual Paulista (UNESP), 1980. Docente no Programa de Pós-Graduação em Ciência da Informação da Universidade Estadual Paulista (PPGCI/UNESP).

Rosa Estopà Bagot (UPF, Espanha)
Doctora en Lingüística por la Universitat Pompeu Fabra (UPF), 1999. Postgraduada en Patología del Lenguaje por la Universitat Autònoma de Barcelona (UAB), 1993. Licenciada en Filología Catalana por la Universitat de Barcelona (UB), 1992. Profesora titular del Departamento de Traducción y Ciencias del Lenguaje de la Universidad Pompeu Fabra (UPF).

Sandra de Albuquerque Siebra (UFPE, Brasil)
Doutora em Ciência da Computação pela Universidade Federal de Pernambuco (2007). Mestre em Ciência da Computação pela Universidade Federal de Pernambuco (1998). Graduada em Ciência da Computação pela Universidade Federal da Paraíba (1995). Professora Adjunta no Departamento de Ciência da Informação da Universidade Federal de Pernambuco (DCI/UFPE). Docente do Programa de Pós-graduação em Ciência da Informação da Universidade Federal de Pernambuco (PPGCI/UFPE).

Sandra Milena Roa Martínez (Unicauca, Colombia)
Doutoranda no Programa de Pós-graduação em Ciência de Informação da Universidade Estadual Paulista (PPGCI/UNESP). Mestre em Engenheira com ênfases em Eletrônica pela Universidad del Valle – Colômbia, 2010. Especialista em Redes de Comunicação da Universidad del Valle – Colômbia, 2004. Graduada em Engenheira de Sistemas da Universidad Industrial de Santander – Colômbia, 1999. Professora Titular do Departamento de Sistemas da Faculdade de Engenharia Eletrônica e Telecomunicações da Universidade do Cauca (Colômbia).

Silvana Aparecida Borsetti Gregorio Vidotti (UNESP, Brasil)
Doutora em Educação pela Universidade Estadual Paulista (UNESP), 2001. Mestre em Ciências - área Ciências da Computação e Matemática Computacional – pela Universidade de São (USP), 1993. Especialista em Ciência da Computação pela Universidade de São (USP), 1987. Licenciada em Matemática pela Universidade Estadual Paulista (UNESP), 1986. Professora Assistente-Doutora em Regime de Dedicação Integral à Docência e à Pesquisa do Departamento de Ciência

da Informação da Universidade Estadual Paulista (DCI/UNESP). Docente do Programa de Pós-graduação em Ciência da Informação da Universidade Estadual Paulista (PPGCI/UNESP).

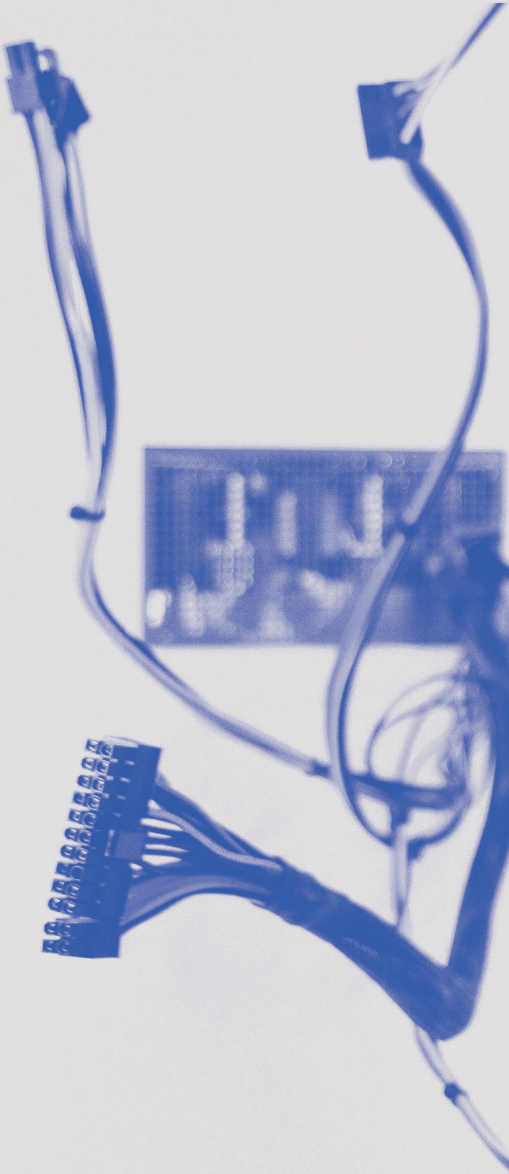
Silvana Drumond Monteiro (UEL, Brasil)
Pós-doutora em Ciência da Informação pela Universidade Federal de Minas Gerais (UFMG), 2014. Doutora em Comunicação e Semiótica pela Pontifícia Universidade Católica de São Paulo, PUC-SP, 2003. Mestre em Ciência da Informação pela Pontifícia Universidade Católica de Campinas (PUCCamp), 1996. Graduada em Biblioteconomia pela Universidade Estadual de Londrina, UEL, 1984. Professora Associada do Departamento de Ciência da Informação da Universidade Estadual de Londrina (DCI/UEL). Docente do Programa de Pós-Graduação em Ciência da Informação da Universidade Estadual de Londrina (PPGCI/UEL).

Vildeane da Rocha Borba (UFPE, Brasil)
Doutoranda em Comunicação e Informação na Universidade Federal do Rio Grande do Sul (UFRGS). Mestre em Ciência da Informação pela Universidade Federal da Paraíba, (UFPB), 2009. Graduada em Biblioteconomia pela Universidade Federal de Pernambuco, (UFPE), 2006. Professora do Departamento de Ciência da Informação da Universidade Federal de Pernambuco (DCI/UFPE).

Este livro foi diagramado pela Editora UFPB em 2018, utilizando a fonte Chaparral Pro. Impresso em papel Offset 75 g/m² e capa em papel Supremo 250 g/m².

Uma fronteira pode ser compreendida como uma linha imaginária que cria uma divisão ou uma delimitação. Nos contextos científicos, as fronteiras delimitam espaços entre o conhecido e o desconhecido, criam territórios de pesquisa e investigação, separam e ao mesmo tempo reúnem pesquisadores e proporcionam o crescimento de pesquisas sobre temáticas específicas.

Nesta obra, refletimos sobre algumas fronteiras científicas relacionadas a “Informação e Tecnologia”, no contexto da Ciência da Informação. Nosso ponto de partida foi a percepção de que as elas são dinâmicas e podem se modificar no processo de acumulação científica. Também nos embasamos no caráter interdisciplinar da Ciência da Informação, que, em nossa concepção, propulsiona novos olhares sobre o conhecido e uma gama de possíveis olhares sobre o desconhecido.



ISBN 978-85-237-1328-7