



Tendências da Pesquisa  
Brasileira em  
Ciência da Informação

## ENLACE DE OBJETOS DIGITAIS NO CONTEXTO DA PUBLICAÇÃO CIENTÍFICA AMPLIADA

*LINKING OF DIGITAL OBJECTS IN THE CONTEXT OF ENHANCED PUBLICATIONS*

Emanuelle Torino<sup>1</sup>

Silvana Aparecida Borsetti Gregorio Vidottiz

Rachel Cristina Vesu Alves<sup>3</sup>

Plácida Leopoldina Ventura Amorim da Costa Santos<sup>4</sup>

**RESUMO:** Vivenciamos, nas últimas décadas, diversas alterações no cenário científico, que trouxeram avanços no seu processo de comunicação com o advento de novos sistemas de informação e na sua forma de disseminação, privilegiando o acesso aberto e o compartilhamento de diferentes objetos gerados no contexto da pesquisa, agregando ao tradicional artigo científico novos produtos, como dados de pesquisa, *slides*, apresentações e vídeos. O enlace desses objetos em torno de uma única pesquisa é entendido na literatura como publicação científica ampliada, que possui diferentes modelos de dados. Nesse sentido, o presente estudo objetiva discutir como enlaçar diferentes objetos digitais no contexto da publicação científica ampliada. O estudo foi realizado a partir de pesquisa bibliográfica para embasamento teórico-conceitual e da pesquisa exploratória para compreensão do cenário abordado. Apresenta o Sistema de Informação de Pesquisa Corrente (CRIS) como infraestrutura adequada para uma ecologia informacional capaz de enlaçar os objetos e seus agentes (pessoas, instituições e agências de fomento), no contexto da publicação científica ampliada. A estruturação do ambiente para que as ligações ocorram de forma satisfatória, seja com os objetos disponibilizados no CRIS ou em qualquer outro ambiente informacional digital, requer conhecimentos de metadados, representação e arquitetura da informação, visando que a camada de dados seja adequada para a recuperação por aplicações computacionais e a interface propícia para a encontrabilidade da informação por humanos.

**Palavras-chave:** Publicação científica ampliada. Metadados. Sistema de informação de pesquisa corrente. CRIS.

---

<sup>1</sup> Universidade Tecnológica Federal do Paraná (UTFPR); Universidade Estadual Paulista – Unesp, [emanuelle@utfpr.edu.br](mailto:emanuelle@utfpr.edu.br), <https://orcid.org/0000-0002-3791-9884>

<sup>2</sup> Universidade Estadual Paulista – Unesp, [silvana.vidotti@unesp.br](mailto:silvana.vidotti@unesp.br), <https://orcid.org/0000-0002-4216-0374>

<sup>3</sup> Universidade Estadual Paulista – Unesp, [rachel.vesu@unesp.br](mailto:rachel.vesu@unesp.br), <https://orcid.org/0000-0002-1649-3722>

<sup>4</sup> Universidade Estadual Paulista – Unesp, [placidasantos@gmail.com](mailto:placidasantos@gmail.com), <https://orcid.org/0000-0003-4478-5623>

**ABSTRACT:** In the last decades, we have experienced several changes in the scientific scenario, which brought advances in its communication process with the advent of new information systems and its dissemination, favoring open access and the sharing of different objects generated in the research context. This fact added to the traditional scientific article new products such as research data, slides, presentations and videos. The binding of these objects around a single research is understood in the specific literature as an enhanced publication, which has different data models. In this sense, the present study aims to discuss how to link different digital objects in the context of enhanced publication. The study was conducted from bibliographic research, for theoretical and conceptual basis, and exploratory research, in order to understand the addressed scenario. It presents the Current Research Information System (CRIS) as an adequate infrastructure for an informational ecology capable of linking objects and their agents (people, institutions and funding agencies), in the context of the enhanced publication. Structuring the environment for connections to occur satisfactorily, whether with objects made available in CRIS or any other digital informational environment, requires knowledge of metadata, representation, and information architecture, so that the data layer is suitable for retrieval by computational applications and the proper interface for the information findability by humans.

**Keyword:** Enhanced Publication. Metadata. Current Research Information System. CRIS.

## 1 INTRODUÇÃO

A crise dos periódicos, na década de 1980, desencadeou alterações significativas no processo de comunicação científica, sobretudo por meio do movimento de acesso aberto, nascido na comunidade de pesquisadores (BUDAPEST OPEN ACCESS INITIATIVE, 2002, 2012; COSTA, 2006; KURAMOTO, 2008; MUELLER, 2006).

Tal iniciativa inegavelmente gerou impactos e até mesmo rupturas, quer seja no protagonismo dos autores na decisão das fontes de publicação que utilizam para a submissão de seus resultados, quer seja na postura recentemente tomada por instituições de pesquisa e agências de fomento, no que tange às negociações com grandes potências editoriais, ou ainda no apoio à manutenção de repositórios digitais de produção científica e de dados de pesquisa. Este movimento teve início na Comunidade Europeia, fortaleceu-se nas iniciativas do Horizonte 2020 (COMISSÃO EUROPEIA, 2014; RICE; SOUTHALL, 2016) e se expandiu para outros continentes.

No contexto brasileiro, a Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) tem sido a pioneira no reconhecimento da necessidade de investimento em pesquisas que tenham caráter aberto, seja na disponibilização em acesso aberto dos seus resultados em repositórios digitais ou na gestão dos dados gerados pela pesquisa (FAPESP, 2018). Esta é uma discussão recente, que ainda necessita de embasamento teórico-conceitual para apoiar a prática nos ambientes informacionais digitais utilizados para a gestão desses conteúdos.

Desta forma, reforça-se a relevância da comunicação científica, tida por Meadows (1999) como o coração da ciência, considerando o incentivo ao acesso, uso e reuso de resultados de pesquisas e dos dados gerados.

Presencia-se um novo cenário de produção, gerenciamento e disponibilização de diferentes objetos digitais resultantes da atividade de pesquisa, como artigos, livros, dissertações, teses, bem como fotografias, apresentações, áudios, vídeos, softwares, *datasets* e os dados de pesquisa. Com isso, uma nova forma de tratamento informacional passou a ser

requerida nas áreas de Biblioteconomia e Ciência da Informação (CI), habituadas ao tratamento informacional. Neste sentido, Rice e Southall (2016) afirmam que há uma resistência entre os bibliotecários para uma mudança de direção, e os alerta para que as bibliotecas, oportunamente, apropriem-se dessa transformação no cenário científico e do conhecimento baseado em dados.

Considerando que os objetos digitais provenientes da atividade de pesquisa, na maioria das vezes, são publicados e/ou disponibilizados em diferentes ambientes informacionais, como fornecer uma visão ampla do processo de concepção, desenvolvimento, disseminação da pesquisa?

Diante do exposto, o presente estudo objetiva discutir como enlaçar diferentes objetos digitais no contexto da publicação científica ampliada.

Para Bardi e Manghi (2014, p. 240, tradução nossa), “‘Publicações científicas ampliadas’ são comumente entendidas como publicações digitais que consistem de uma parte narrativa obrigatória (a descrição da pesquisa realizada) mais ‘partes’ relacionadas, como *datasets*, outras publicações, imagens, tabelas, *workflows*, planos.”

A metodologia utilizada para a realização deste estudo foi a de pesquisa bibliográfica para embasamento teórico-conceitual e de pesquisa exploratória para compreensão do cenário abordado.

## **2 PUBLICAÇÃO CIENTÍFICA AMPLIADA**

As iniciativas de acesso aberto e o advento das tecnologias de informação e comunicação (TIC) possibilitaram que diferentes ambientes - como periódicos, anais eletrônicos, catálogos online de editoras, repositórios digitais ou ainda SlideShare, Flickr, blogs, páginas pessoais de instituições acadêmicas e de pesquisa - fossem desenvolvidos para que os pesquisadores pudessem publicar e/ou disponibilizar seus conteúdos.

O conceito de liberdade de compartilhamento advém da Ciência da Computação, por meio do movimento de software livre, cujo precursor é Richard Stallman, para o qual a

cooperação é uma ação humana natural, o que o levou a afirmar, no Manifesto GNU (2018), em 1985, que os códigos de softwares devem ser compartilhados para que possam ser melhorados ou adequados às necessidades dos interessados e que o contrário consiste em uma quebra de solidariedade. Isso levou ao conceito de *copyleft* que, contrariamente ao *copyright*, possibilita o compartilhamento e a cooperação.

No campo científico, o acesso aberto trouxe ao autor o entendimento de que ele pode manter a titularidade dos seus direitos autorais e depositar o resultado de sua pesquisa em um repositório digital da instituição com a qual mantém vínculo, ainda que seja necessário respeitar um período de restrição de acesso solicitado pela fonte de publicação. Da mesma forma, acompanha-se uma evolução na disponibilização dos dados, alterando a perspectiva da publicação como mercadoria para o compartilhamento, o que deve ser incentivado pelos bibliotecários (RICE; SOUTHALL, 2016).

De igual maneira, há iniciativas para a abertura de dados de pesquisa, que deve se concretizar em um período menor, considerando as solicitações, sobretudo das agências de fomento. A Comunidade Europeia define que o acesso aberto a dados de pesquisa consiste na possibilidade de reutilização, considerando as especificações legais, com vistas a basear-se em um diálogo científico extenso e avançar no aprimoramento de pesquisas anteriores, incentivar a colaboração, acelerar a inovação e envolver indivíduos e a sociedade. “Em um contexto de pesquisa, os exemplos de dados incluem estatísticas, resultados de experimentos, medições, observações resultantes do trabalho de campo, resultados de pesquisas, gravações de entrevistas e imagens.” (EUROPEAN COMMISSION, 2017).

A base da disponibilização em acesso aberto, seja de softwares, artigos de periódicos, dados de pesquisa ou qualquer objeto gerado na atividade de pesquisa, é a perspectiva de que não deve haver pagamento para o acesso, uso, reuso consciente e qualquer forma de aproveitamento estipulada legalmente por qualquer organismo público ou privado, bem como pela sociedade, uma vez que as pesquisas são, geralmente, desenvolvidas e financiadas pelo erário.

Desta forma, expande-se o horizonte da publicação e/ou disponibilização de objetos resultantes da atividade científica, que certamente devem ser gerenciados por plataformas de softwares com objetivos específicos, que necessitam de estruturas consistentes de representação e enlace, constituindo-se uma rede de objetos científicos, contendo: dados de pesquisa, artigos, fotografias, apresentações, tabelas, *datasets*, softwares, vídeos, dissertações, teses, livros, trabalhos apresentados em eventos, relatórios e outros tipos, ligados entre si por um processo único de pesquisa, do qual são resultantes. Fato que, segundo Van de Sompel e Lagoze (2007), requer mudanças na comunicação científica e no entendimento do que constitui um objeto nesse novo ambiente.

A publicação e o reconhecimento de autoria dos pesquisadores já é consenso no que tange às publicações tradicionais, fazê-lo, considerando o advento das formas de compartilhamento e reuso de publicações ou de dados de pesquisa, ainda é tema a ser discutido e para o qual a Ciência da Informação deve prestar suas contribuições.

A este respeito, reconhecem-se novas formas de publicação em periódicos, a exemplos dos artigos tecnológicos (GREGOR; HEVNER, 2013; MOTTA, 2017) e artigos de dados (RICE; SOUTHALL, 2016; ROA-MARTÍNEZ; VIDOTTI; SANT'ANA, 2017), que podem consistir em uma das vias para que novos produtos de pesquisa também sejam reconhecidos, adequadamente citados e alcancem prestígio em suas áreas. Há de se atentar para a característica de reprodutibilidade da ciência, de forma a reforçar ou refutar seus resultados, e a disponibilização dos dados pode ser benéfica.

Isso requer que os pesquisadores passem a documentar de forma mais específica seus dados de pesquisa, quer seja na descrição em artigos de dados, quer seja na representação, incluindo o seu contexto, por meio de metadados. A este respeito, a Ciência da Informação precisa atuar em diferentes aspectos, como na definição da infraestrutura necessária à disponibilização de diferentes tipos de dados e demais objetos resultantes da atividade de pesquisa, e também na estruturação das diferentes camadas das ferramentas de disponibilização de produtos de pesquisa, tanto na estruturação dos metadados, quanto na sua representação e na interface, para que haja encontrabilidade.

Para Van de Sompel e Lagoze (2007, p. 1, tradução nossa), o ambiente de pesquisa moderno está constituído por unidades compostas, que são “agregações de unidades de informação distintas que, quando combinadas, formam um todo lógico”, que independem do formato digital, da tipologia, do tipo e formato de mídia e podem ser dinâmicas.

Podemos entender que essas agregações de unidades de informação distintas vão compor uma nova forma de publicação científica, a publicação científica ampliada. Para os autores Bardi e Manghi (2014), a publicação científica ampliada, do inglês enhanced publications (EP), consiste em uma parte textual obrigatória para descrever a pesquisa e subpartes enlaçadas. Para a descrição da pesquisa, se estabelecem identificador persistente e metadados, que também podem ser incluídos às suas subpartes.

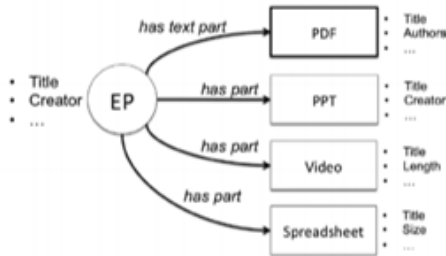
A publicação científica ampliada objetiva que os produtos de pesquisa primários, amplamente reconhecidos (artigos, dissertações, teses, trabalhos publicados em eventos, livros, relatórios), possam ser expandidos com outros produtos, provenientes do mesmo processo de pesquisa, e que eles possam se interconectar de forma legível a humanos e aplicações computacionais (ARAYA, 2014; BARDI e MANGHI, 2014).

Bardi e Manghi (2014) estudaram os modelos de dados de publicação científica ampliada visando identificar seus padrões, sobretudo de metadados e associações e os segmenta em classes, considerando suas partes recorrentes. Assim, de acordo com os autores, podemos encontrar os seguintes modelos (Figura 1) para a publicação científica ampliada:

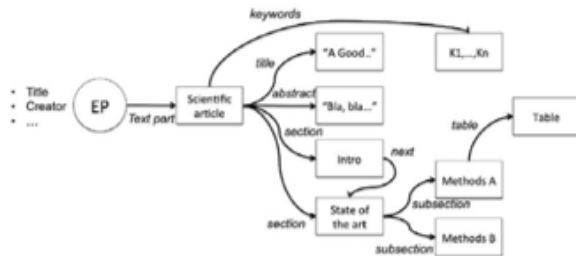
- a) modelo de dados com partes embutidas;
- b) modelo de dados com partes estruturadas de texto;
- c) modelo de dados com partes de referência;
- d) modelo de dados com partes executáveis;
- e) modelo de dados com partes geradas.

Figura 1 - Publicação ampliada: modelos de dados

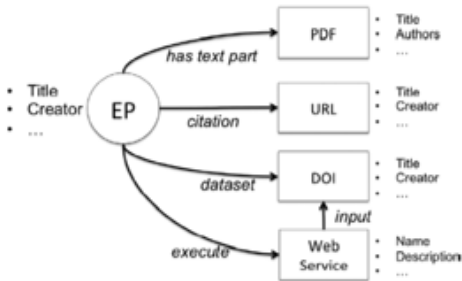
**A** partes embutidas



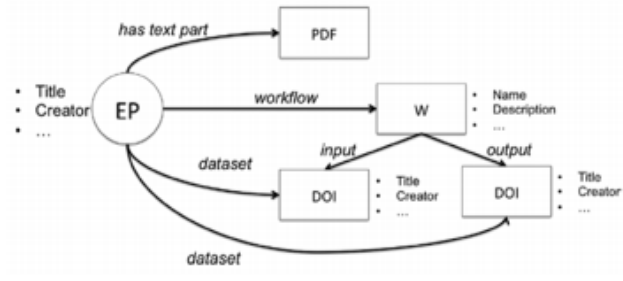
**B** partes estruturadas de texto



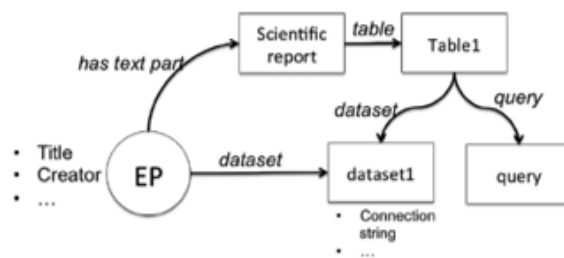
**C** partes de referência



**D** partes executáveis



**E** partes geradas



Fonte: Adaptado de Bardi e Manghi (2014).

No modelo de dados com partes embutidas, ao arquivo textual é adicionado contexto com outros objetos decorrentes da pesquisa, que são disponibilizados em um único registro, como material suplementar. Os objetos podem ser de diferentes formatos e apresentar metadados específicos, conforme a Figura 1 modelo A. Considerando a disponibilização em um único registro, há um identificador e metadados comuns, dificultando a recuperação e o compartilhamento individual dos objetos.



No modelo de dados com partes estruturadas de texto, a noção da publicação com um único arquivo, a exemplo de um PDF, é adaptada para o contexto da web e pretende que o leitor faça uso do material a partir do acesso às suas partes, que são conectadas por metadados, com a possibilidade de gerar um arquivo único com enlace das partes. Neste modelo, representado no modelo B da Figura 1, a recuperação da informação é realizada pelo conteúdo do material.

O modelo de dados com partes de referência possibilita que diferentes objetos sejam disponibilizados, possuam identificadores e metadados específicos e possam ser ligados entre si de modo referencial, conforme modelo C na Figura 1. Para tanto, é necessária a adição de metadados para enlaçar as partes; há sistemas de informação que permitem que tais metadados sejam herdados do material textual. Neste modelo, os objetos podem ser recuperados individualmente e direcionar o usuário a todas as partes que constituem a publicação ampliada, respeitando suas características, formas de armazenamento e gestão. Uma dificuldade neste tipo de enlace são os *links* quebrados, cuja ocorrência pode ser minimizada com a adoção de identificadores persistentes para todas as partes referenciadas.

O modelo de dados com partes executáveis, modelo D da Figura 1, está apoiado na perspectiva de reprodutibilidade científica, de forma que o objeto textual gerado está acompanhado de um fluxo de atividade, os dados que o geraram, os que foram gerados por ele e um conjunto de metadados capaz de descrever adequadamente todo o contexto, incluindo processos e calibragem de equipamentos, para que a pesquisa possa ser reproduzida e os resultados comparados.

O modelo de dados com partes geradas, modelo E da Figura 1, diferente dos demais, trabalha com partes dinamicamente geradas por aplicações, cujos parâmetros e configurações são estabelecidos no sistema e descritos nos metadados, mantendo as partes geradas intimamente ligadas e dependentes do sistema.

Araya (2014) descreve os sistemas de informação utilizados à época para gerenciar produtos de pesquisa, contudo destaca que a infraestrutura ainda se concentrava no armazenamento e disseminação de objetos individuais sem realizar o enlace necessário aos

objetos compostos e afirma que estes enlaces necessitam ser compreendidos por humanos e aplicações computacionais. Desta forma, esclarece que a publicação científica ampliada necessita de suporte técnico e funcional para administração e armazenamento.

Isto posto, retornamos ao tratamento informacional, base da Ciência da Informação, que atua, dentre outros, em processos relacionados à representação, ao armazenamento, ao acesso e à disseminação da informação, tradicionalmente aplicados a objetos textuais. No contexto da publicação científica ampliada, há, ainda, a necessidade do tratamento informacional de objetos não-textuais e, mais recentemente, de dados de pesquisa.

Vale destacar que os processos realizados pela CI, no que tange à informação, consistem na estrutura de sistemas e armazenamento de dados que representam conteúdos, para que sejam recuperáveis pelo usuário.

Com o advento das TIC, a área se vale de técnicas e tecnologias para tratamento e representação da informação, descrevendo-a em dados, que mais recentemente são gerenciados e manipulados por aplicações computacionais, por meio dos metadados.

### **3 ENLACE DE OBJETOS PARA A PUBLICAÇÃO CIENTÍFICA AMPLIADA**

Os modelos de dados apresentados por Bardi e Manghi (2014) elucidam as formas de produção, comunicação e disponibilização de objetos na publicação científica ampliada. O primeiro modelo, de dados com partes embutidas, consiste na forma tradicional de publicação de artigos em periódicos digitais, nos quais, ao registro texto completo, pode(m) ser incluído(s) material(is) suplementar(es), em sua maioria como outro(s) arquivo(s), sem qualquer representação ou possibilidade de recuperação e compartilhamento individual.

Os avanços dessas mesmas plataformas permitiram a publicação de trabalhos em outros formatos, que beneficiam as operações de aplicações computacionais, como os arquivos *HyperText Markup Language* (HTML) e *Extensible Markup Language* (XML). Independente do formato, nessas plataformas também é possível inserir material suplementar, reforçando o modelo com partes embutidas.

Vale destacar as ponderações de Van de Sompel e Lagoze (2007) acerca da representação e disponibilização de objeto composto, caracterizado pela agregação de diferentes recursos informacionais em uma unidade de publicação científica como que constituindo um “todo”, para o qual se atribui um identificador, que pode ser interpretado por humanos, mas que prejudica a recuperação por agentes computacionais, visto que esses interpretam a página principal como sendo de um recurso e o adequado seria possibilitar a interpretação de um contexto que relaciona diferentes recursos, constituindo, assim, um objeto composto e permitindo, ainda, que um único recurso possa ser parte de diferentes objetos compostos.

O advento de novos softwares de apoio à atividade de pesquisa privilegiou a implantação e a disponibilização de ambientes informacionais que permitem a gestão de diferentes objetos de pesquisa, que podem ser referenciados, constituindo uma rede de objetos acessíveis em múltiplas fontes. Neste caso é possível, por exemplo, enlaçar um artigo disponível em periódico com os dados de pesquisa disponíveis em repositório de dados, e ainda com vídeo, apresentações em slides, trabalho apresentado em evento ou mesmo uma tese, disponíveis em um repositório digital.

No exemplo citado, o enlace pode ser realizado por meio de metadados de relacionamento, sendo possível utilizar os relacionamentos em uma via, ligando os ambientes internos a uma instituição aos ambientes externos a ela e em duas vias apenas em ambientes internos que compõem uma ecologia informacional.

Vale considerar que, no caso do relacionamento em uma via, cada instituição é detentora do seu ciclo de vida de dados e compete apenas a ela a prerrogativa de realizar o enlace com objetos disponíveis em fonte(s) externa(s). A dificuldade, neste sentido, consiste em enlaçar objetos já disponíveis quando um novo objeto relacionado ao mesmo processo de pesquisa é gerado e disponibilizado, o que se deve à impossibilidade de atuação do produtor dos recursos nos diferentes ambientes em que seus produtos de pesquisa são disponibilizados. Esta dificuldade, que consiste na gestão dos ambientes digitais, dificulta o enlace dos objetos e impede que o usuário da informação navegue por todos aqueles objetos

que foram gerados a partir de uma única pesquisa, isso posto, verifica-se que o relacionamento pode parecer incompleto.

Os ambientes digitais devem ser amigáveis para o usuário; possibilitar a ligação com outros dados, ainda que por *links*; utilizar identificadores persistentes e favorecer a recuperação por agentes computacionais, o que reforça a relevância da adoção dos princípios FAIR (*findable, accessible, interoperable, reusable*), para que os registros sejam encontráveis, acessíveis, interoperáveis e reutilizáveis (EUROPEAN COMMISSION, 2016).

Os princípios FAIR se referem ao processamento dos agentes computacionais para localizar, acessar, interoperar e reutilizar dados com a mínima intervenção humana. Tais princípios se referem a três entidades: dados ou qualquer objeto digital, metadados e infraestrutura (GOFAIR, 2019).

A partir da análise de dados de publicação científica ampliada realizada por Bardi e Manghi, (2014, p. 243, tradução nossa)

Para descrições de partes de metadados, observamos que eles fornecem informações em diferentes níveis de interpretação, de modo a permitir a interpretação tanto humana quanto mecânica [...] informações bibliográficas, informações de arquivos, informações de proveniência, informações de visualização, informações de execução, informações de versão.

Os dados de pesquisa e demais objetos por ela gerados devem fornecer metadados bibliográficos que identifiquem a publicação (EUROPEAN COMMISSION, 2017). Entendemos ser necessário expandir a compreensão de que a estes devem ser somados os metadados técnicos, administrativos, de uso, de preservação e de proveniência.

Bardi e Manghi (2014), em seu estudo, buscaram compreender as características estruturais e semânticas das publicações ampliadas e verificaram que cada parte pode conter descrições próprias de metadados para assegurar diferentes usos, por exemplo, metadados bibliográficos, de proveniência, formato de arquivo, informações de visualização e de versão. Exemplificam ainda que existem informações de relacionamento entre as partes que podem se constituir em orientações aos usuários (por exemplo: *chapterOf, relatedWith, datasetUsed*) ou a aplicações (por exemplo: *alternativeVisualization, externalLink, localLink*).

Este é ainda um desafio para a representação da informação, ao passo que antecede dois outros modelos apresentados por Bardi e Manghi (2014), o de partes executáveis e o de partes geradas, esses últimos trazem a dinamicidade para os dados gerados, característica esta diferente da representação de objetos finalizados.

As novas possibilidades de publicação científica, têm trazido a representação da informação para a figura do autor enquanto responsável pela inserção dos seus conteúdos em diferentes ambientes informacionais. Destaca-se que a representação é uma atividade que exige técnicas específicas, para a qual a maioria dos autores não está capacitada. Por outro lado, a necessidade de agilidade no processo, bem como a falta de clareza dos impactos negativos advindos do processamento inadequado para a recuperação da informação, tem afastado os profissionais da CI deste cenário.

Para Araya (2014), é crescente a quantidade de objetos informacionais disponíveis no ciberespaço, a limitação está em encontrá-los e relacioná-los às pesquisas das quais fazem parte, ao que corroboram Van de Sompel e Lagoze (2007, p. 6, tradução nossa), “A reutilização de um recurso da Web depende da capacidade de fazer referência a ele”. Araya (2014) menciona ainda a necessidade de “uma infraestrutura que possibilitasse a efetiva vinculação dos objetos compostos e o consequente enriquecimento da publicação científica” (p. 95), capaz de operacionalizar a publicação científica ampliada.

Nesse cenário, destaca-se o *Current Research Information System* (CRIS) (ASSERSON; JEFFERY, 2009; JEFFERY; ASSERSON, 2009; LOPATENKO, 2001; 2010; SHEPPARD, 2010), cuja tradução é sistema de informação de pesquisa corrente, que pode se constituir em uma plataforma capaz de gerenciar o ciclo de vida da pesquisa, desde a sua concepção até a conclusão, e ainda interligar os resultados gerados, os agentes envolvidos (pessoas, instituições e agências de fomento) e seus relacionamentos, de forma a atuar como ferramenta de apoio à tomada de decisão por parte de instituições e gestores de pesquisa.

Muitas dessas informações já estão disponíveis em sistemas de informação complementares com funcionamento independente – a exemplo do Diretório de Instituições, Diretório de Grupos de Pesquisa, Currículo Lattes, Bibliotecas Virtuais de Agências de Fomento

– e que podem, uma vez integrados, oferecer uma infraestrutura de informação completa sobre a atividade de pesquisa, no contexto de uma instituição em primeira instância, e até mesmo como federações nacionais.

Para que isso ocorra, é necessário que os sistemas de informação sejam adequadamente integrados, considerando suas características e contextos.

Questionado sobre o que é um CRIS, Simon Kerridge afirma que há alguma complexidade em oferecer uma definição, mas que pode ser entendido como a utilização de tecnologia, especialmente informática, para a melhoria dos processos administrativos da pesquisa, essencialmente para gerenciar informação de pesquisa em vários aspectos: pessoas (pesquisadores); publicações (resultados ou produtos de pesquisa); projetos e propostas (informações sobre financiamento); informações sobre a pós-graduação; impacto; ética em pesquisa; indicadores de desempenho (SHEPPARD, 2010).

De forma muito assertiva, Joint (2008, p. 571, tradução nossa) define CRIS como um:

[...] sistema de informação que pode gerenciar toda a informação relevante da pesquisa, começando com oportunidades de financiamento, passando pelo estágio de redação e submissão de propostas, seguindo com as propostas bem-sucedidas, que se tornam projetos ativos que serão gerenciados até a conclusão – estágio no qual são gerados resultados, muitos dos quais são publicações ou algum outro artefato da atividade de pesquisa.

Vale destacar que nessa estrutura o CRIS ocupa o centro do processo, liga diferentes sistemas e bancos de dados e torna-se a plataforma central de atividade de pesquisadores e gestores de pesquisa (ASSERSON; JEFFERY, 2009).

Para tanto, no contexto Europeu, há um modelo de referência para o desenvolvimento de CRIS, tido como recomendação oficial, o *Common European Research Information Format* (CERIF), mantido pelo euroCRIS como um padrão internacional de dados relacionais para o armazenamento e interoperabilidade de informação de pesquisa, que trabalha com o modelo entidade-relacionamento.

“CERIF desponta como a norma mais importante e conceitualmente mais bem elaborada [...] [para] o intercâmbio de informações entre as plataformas CRIS no âmbito da Comunidade Européia.” (SALES; SAYÃO, 2015, p. 170).

O modelo relacional CERIF, de acordo com Sales e Sayão (2015), define quatro níveis de entidades: básicas, resultados, segundo nível e *links*. Esse modelo tem como foco e entidade principal o projeto, que, juntamente com organização e pessoa, configura o nível básico e o núcleo do modelo, do qual também fazem parte entidades que expressam seus desdobramentos: resultado (publicação, patente, produto). No segundo nível estão as entidades que ‘capacitam’ o modelo e representam o contexto da pesquisa. A entidade *link* estabelece os relacionamentos, por isso é considerada pelo euroCRIS o ponto forte do modelo, pode ser vinculada a diferentes elementos – pessoa e publicação, pessoa e organização, organização e publicação, projeto, pessoa e publicação – e garante a dinâmica do contexto científico.

Desta forma, o CRIS pode se constituir em uma ecologia informacional institucional de modo a enlaçar as pesquisas nela realizadas com os dados de pesquisa por ela gerados e os objetos resultantes dessas pesquisas, estabelecendo ligações por meio do CERIF e dos metadados de relacionamento.

#### **4 CONSIDERAÇÕES FINAIS**

Diferentes alterações, quer sejam decorrentes do advento das TIC que oportunizaram o desenvolvimento de ambientes informacionais digitais, quer sejam das mudanças no processo de comunicação científica, favoreceram a disponibilização de objetos informacionais resultantes de atividades de pesquisa que precisam estabelecer um enlace de forma a representá-los individualmente e, paralelamente, torná-los parte de um todo, considerando a publicação científica ampliada.

O CRIS tem se desenvolvido como ambiente propício para a disponibilização e gestão de informações e objetos provenientes das atividades de pesquisa, apoiado no modelo CERIF,

que estabelece um padrão entidade-relacionamento para a descrição e enlace desses objetos no ambiente.

A estruturação do ambiente para que as ligações ocorram de forma satisfatória, seja com os objetos disponibilizados no CRIS ou em qualquer outro ambiente informacional digital, requer conhecimentos de metadados, representação e arquitetura da informação, visando que a camada de dados seja adequada para a recuperação por aplicações computacionais e a interface benéfica para a encontrabilidade por humanos.

Nesse contexto, o profissional da informação, especificamente o bibliotecário, deve valer-se dos conhecimentos advindos da estrutura base da sua formação para a atuação, de forma a beneficiar pesquisadores, usuários e instituições de pesquisa e de fomento no sentido de tornar os objetos resultantes da atividade de pesquisa disponíveis de forma aberta, atentando-se aos aspectos legais, para que possam atingir os objetivos da publicação científica ampliada e do acesso aberto: de tornar os objetos resultantes das atividades de pesquisa acessíveis, recuperáveis, interoperáveis e reutilizáveis pela comunidade.

## REFERÊNCIAS

ARAYA, E. R. M. **Comunicação científica**: agregação, compartilhamento e reuso de elementos informacionais. 2014. 130 f. Tese (doutorado) - Universidade Estadual Paulista “Júlio de Mesquita Filho”, 2014. Disponível em: <http://hdl.handle.net/11449/121981>. Acesso em: 13 dez. 2018.

ASSERSON, A.; JEFFERY, K. G. Current research information systems (CRIS): past, present and future. **Wissenschaftsmanagement**, n. 1, p. 41-45, jan./feb. 2009. Disponível em: <http://bora.uib.no/handle/1956/6929>. Acesso em: 16 dez. 2018.

BARDI, A.; MANGHI, P. Enhanced publications: data models and information systems. **LIBER Quarterly**, [S.l.], v. 23, n. 4, p. 240-273, abr. 2014. ISSN 2213-056X. Disponível em: <https://dspace.library.uu.nl/handle/1874/292632>. Acesso em: 13 dez. 2018.



BUDAPEST OPEN ACCESS INITIATIVE. **Dez anos da iniciativa de Budapeste em acesso aberto:** a abertura como caminho a seguir. 2012. Disponível em:

<https://www.budapestopenaccessinitiative.org/boai-10-translations/portuguese-brazilian-translation>. Acesso em: 01 nov. 2018.

BUDAPEST OPEN ACCESS INITIATIVE. **Iniciativa de Budapeste pelo acesso aberto.** 2002.

Disponível em: <https://www.budapestopenaccessinitiative.org/translations/portuguese-translation>. Acesso em: 01 nov. 2018.

COMISSÃO EUROPEIA. **Horizon 2020 em breves palavras:** o programa-quadro de investigação e inovação da UE. 2014. Disponível em:

[http://ec.europa.eu/programmes/horizon2020/sites/horizon2020/files/H2020\\_PT\\_KI0213413PTN.pdf](http://ec.europa.eu/programmes/horizon2020/sites/horizon2020/files/H2020_PT_KI0213413PTN.pdf). Acesso em: 13 out. 2018.

COSTA, S. M. S. Filosofia aberta, modelos de negócios e agências de fomento: elementos essenciais a uma discussão sobre o acesso aberto à informação científica. **Ciência da Informação**, Brasília, v. 35, n. 2, p. 39-50, maio/ago. 2006. Disponível em:

<http://revista.ibict.br/ciinf/article/view/1139/1295>. Acesso em: 13 dez. 2018.

EUROPEAN COMMISSION. **H2020 Programme:** guidelines on FAIR data management in Horizon 2020. 2016. Disponível em:

[http://ec.europa.eu/research/participants/data/ref/h2020/grants\\_manual/hi/oa\\_pilot/h2020-hi-oa-data-mgt\\_en.pdf](http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf). Acesso em: 13 dez. 2018.

EUROPEAN COMMISSION. **H2020 Programme:** guidelines to the rules on open access to scientific publications and open Access to research data in Horizon 2020 . 2017. Disponível em:

[http://ec.europa.eu/research/participants/data/ref/h2020/grants\\_manual/hi/oa\\_pilot/h2020-hi-oa-pilot-guide\\_en.pdf](http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-pilot-guide_en.pdf). Acesso em: 13 dez. 2018.

FAPESP. **Plano de Gestão de Dados FAPESP.** Disponível em:

<http://www.fapesp.br/gestaodedados/>. Acesso em: 13 dez. 2018.

GOFAIR. **FAIR principles.** Disponível em: <https://www.go-fair.org/fair-principles/>. Acesso em: 19 jul. 2019.

GREGOR, S.; HEVNER, A. R. Positioning and presenting design science research for maximum impact. **MIS Quarterly**, v. 37, n. 2, p. 337-356, 2013. Disponível em:

<http://dl.acm.org/citation.cfm?id=2535658.2535660>. Acesso em: 13 dez. 2018.

JEFFERY, K. G.; ASSERSON, A. Institutional Repositories and Current Research Information Systems. **New Review of Information Networking**, v. 14, n. 2, p. 71-83, 2009. Disponível em: <http://dx.doi.org/10.1080/13614570903359357>. Acesso em: 16 dez. 2018.

JOINT, N. Current research information systems, open access repositories and libraries: ANTAEUS. **Library Review**, v. 57, n. 8, p. 570-575, 2008. Disponível em: <https://doi.org/10.1108/00242530810899559>. Acesso em: 19 dez. 2017.

KURAMOTO, H. Acesso livre: um caso de soberania nacional? In: COLÓQUIO MEDIAÇÕES E USOS DE SABERES E INFORMAÇÃO, 1., 2008, Rio de Janeiro. **Anais...** Rio de Janeiro, 2008. p. 4-7.

LOPATENKO, A. S. **Information retrieval in Current Research Information Systems**. 2001. Disponível em: <https://arxiv.org/abs/cs/0110026>. Acesso em: 16 dez. 2018.

MANIFESTO GNU. Disponível em: <http://www.gnu.org/gnu/manifesto.html>. Acesso em: 13 dez. 2018.

MEADOWS, A. J. **A comunicação científica**. Brasília: Briquet de Lemos, 1999.

MOTTA, G. da S. Como escrever um bom artigo tecnológico? **Revista de Administração Contemporânea**, Curitiba, v. 21, n. 5, p. 4-8, out. 2017. Disponível em: [http://www.scielo.br/scielo.php?script=sci\\_arttext&pid=S1415-65552017000500004&lng=en&nrm=iso](http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1415-65552017000500004&lng=en&nrm=iso) >. Acesso em: 13 dez. 2018.

MUELLER, S. P. M. A comunicação científica e o movimento de acesso livre ao conhecimento. **Ciência da Informação**, Brasília, v. 35, n. 2, p. 27-38, maio/ago. 2006. Disponível em: <http://revista.ibict.br/ciinf/article/view/1138/1293>. Acesso em: 13 dez. 2018.

RICE, R.; SOUTHALL, J. **The data librarian's handbook**. London: Facet Publishing, 2016. cap. 10, p. 147-159.

ROA-MARTÍNEZ, S. M.; VIDOTTI, S. A. B. G.; SANT'ANA, R. C. Estructura propuesta del artículo de datos como publicación científica. **Revista Española de Documentación Científica**, v. 40, n. 1, enero-marzo 2017. Disponível em: <http://redc.revistas.csic.es/index.php/redc/article/view/971>. Acesso em: 16 dez. 2018.

SALES, Luana Farias; SAYÃO, Luís Fernando. Ciberinfraestrutura de informação para a pesquisa: uma proposta de arquitetura para a integração de repositórios e sistemas CRIS. **Inf. & Soc.: Est.**, João Pessoa, v. 25, n. 3, p. 163-184, set./dez. 2015. <<http://www.ies.ufpb.br/ojs/index.php/ies/article/view/23998/14535>>. Acesso em 11 ago. 2017.

SHEPPARD, Nick. Learning how to play nicely: repositories and CRIS. **Ariadne**, n. 64, jul. 2010. Disponível em: <http://www.ariadne.ac.uk/issue64/wrn-repos-2010-05-rpt>. Acesso em: 16 dez. 2018.

VAN DE SOMPEL, H., LAGOZE, C. Interoperability for the discovery, use, and re-use of units of scholarly communication. **CTWatch Quarterly**, v. 3, n. 3, ago. 2007. Disponível em: <http://www.ctwatch.org/quarterly/articles/2007/08/interoperability-for-the-discovery-use-and-re-use-of-units-of-scholarly-communication/>. Acesso em: 16 dez. 2018.