

UNIVERSIDADE FEDERAL DE SÃO CARLOS  
CENTRO DE EDUCAÇÃO E CIÊNCIAS HUMANAS  
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO

DÉBORA MARROCO NININ

*LINKED OPEN DATA EM COLEÇÕES DE PATRIMÔNIO CULTURAL: ASPECTOS DA  
REPRESENTAÇÃO DA INFORMAÇÃO PARA HUMANIDADES DIGITAIS*

*LINKED OPEN DATA IN CULTURAL HERITAGE COLLECTIONS: ASPECTS OF  
INFORMATION REPRESENTATION FOR DIGITAL HUMANITIES*

São Carlos/SP  
2018

DÉBORA MARROCO NININ

*LINKED OPEN DATA EM COLEÇÕES DE PATRIMÔNIO CULTURAL: ASPECTOS DA  
REPRESENTAÇÃO DA INFORMAÇÃO PARA HUMANIDADES DIGITAIS*

*LINKED OPEN DATA IN CULTURAL HERITAGE COLLECTIONS: ASPECTS OF  
INFORMATION REPRESENTATION FOR DIGITAL HUMANITIES*

Dissertação apresentada à banca de defesa de Mestrado do Programa de Pós-graduação em Ciência da Informação da Universidade Federal de São Carlos, como requisito para obtenção do título de mestre em Ciência da Informação.

**Área de Concentração:** Conhecimento, Tecnologia e Inovação

**Linha de Pesquisa:** Tecnologia, Informação e Representação

**Orientador(a):** Profa. Dra. Ana Carolina Simionato



grupo de pesquisa  
dados e metadados

São Carlos/SP  
2018



UNIVERSIDADE FEDERAL DE SÃO CARLOS

Centro de Educação e Ciências Humanas  
Programa de Pós-Graduação em Ciência da Informação

---

**Folha de Aprovação**

---

Assinaturas dos membros da comissão examinadora que avaliou e aprovou a Defesa de Dissertação de Mestrado da candidata Débora Marroco Ninin, realizada em 30/08/2018:

---

Profa. Dra. Ana Carolina Simionato  
UFSCar

---

Prof. Dr. José Eduardo Santarem Segundo  
UNESP

---

Prof. Dr. Rogério Aparecido Sá Ramalho  
UFSCar

Certifico que a defesa realizou-se com a participação à distância do(s) membro(s) José Eduardo Santarem Segundo, Rogério Aparecido Sá Ramalho e, depois das arguições e deliberações realizadas, o(s) participante(s) à distância está(ão) de acordo com o conteúdo do parecer da banca examinadora redigido neste relatório de defesa.

---

Profa. Dra. Ana Carolina Simionato

## AGRADECIMENTOS

São muitas as pessoas que precisam ser lembradas nesse momento.

Agradeço primeiramente a Deus, pela vida e toda a sua beleza. Por permitir que eu tenha realizado esse trabalho.

Ao José, pelo infinito amor, carinho e apoio em todos os momentos.

Aos meus pais Roberta e Acácio, à minha irmã Júlia e aos meus avós Sônia e Geraldo, pelo carinho, paciência e, principalmente, pelo exemplo de vida de cada um.

Ao Sebastian, por sua amizade e conselhos acadêmicos sempre bem-vindos.

À Nádea, também pelos conselhos e momentos de descontração.

À Maria Gabriela, Tamiris, Camila, Mariana e Joaquim, simplesmente por serem amigos e estarem sempre prontos para uma palavra de incentivo e para muitas risadas.

À minha orientadora, Ana Carolina, por todo o apoio, paciência e principalmente, pela confiança depositada em mim ao longo desse processo.

Aos colegas do PPGCI Raquel, Júlio, Paulo, Humberto, Guilherme, Melina, Laila, Tamie e Jéssica, pela companhia e apoio mútuo ao longo desse período. Tenham certeza de que nossos almoços nunca serão esquecidos.

Aos professores do PPGCI, por toda a dedicação, esforço e empenho para que o sonho do Programa se realizasse.

Aos membros da banca de qualificação e defesa, José Eduardo e Rogério, pela dedicação e preciosas sugestões.

À Câmara Municipal de São Carlos por permitir e apoiar a realização desse trabalho.

Muito obrigada!

"[...] grandes inovadores são aqueles que conectam arte  
com ciência, humanidades com tecnologia"  
Walter Issacson, sobre Leonardo da Vinci

## RESUMO

**Introdução:** O advento da Internet e da plataforma *Web* suscitam novas possibilidades e desafios de gerenciamento e acesso ao conhecimento, armazenado nos ambientes digitais na forma de dados. Nesse contexto, destacam-se as ferramentas de descrição do movimento *Linked Open Data*, por meio dos princípios *Linked Data*. Em se tratando de dados de patrimônio cultural, presentes em bibliotecas, arquivos e museus, a conexão e disponibilização desses tipos de dados oferecem novas possibilidades de aprendizado e de divulgação dos conhecimentos em Humanidades e do vasto patrimônio cultural distribuído pelo globo. Diante desse cenário, questiona-se: como os princípios *Linked Data* contribuem para a representação e o acesso às coleções de patrimônio cultural? **Objetivo:** Analisar o *Linked Open Data* em instituições de patrimônio cultural, no intuito de contribuir para os estudos acerca da representação e acesso às coleções de patrimônio cultural. **Procedimentos metodológicos:** Trata-se de uma pesquisa de natureza teórico-aplicada, com caráter exploratório e abordagem qualitativa, realizada por meio de levantamento bibliográfico e documental acerca da questão de pesquisa. A bibliografia científica e os documentos pertinentes foram analisados de modo a fornecer subsídios para a representação e promoção do acesso às coleções de patrimônio cultural no escopo do movimento *Linked Open Data*. **Resultados:** Os resultados alcançados envolveram 1) contribuições acadêmicas a respeito dos benefícios do uso das ferramentas do movimento *Linked Open Data*, destacando-se questões de interoperabilidade (técnica e semântica) e de sustentabilidade de dados, 2) levantamento de iniciativas *Linked Open Data* em instituições de patrimônio cultural, 3) apresentação e discussão dos modelos de dados adotados pelas iniciativas Europeia e DPLA e do esquema de dados do projeto *DigitalNZ* e 4) apresentação das soluções desenvolvidas pela comunidade de patrimônio cultural para uso dos padrões de metadados para a descrição *Linked Open Data* de coleções de patrimônio cultural. Estes resultados demonstraram que a comunidade de patrimônio cultural tem buscado se aproveitar dos benefícios oferecidos pelo movimento *Linked Open Data*, especialmente no que diz respeito à sustentabilidade dos processos de descrição e acesso aos dados de coleções de patrimônio cultural e que, com isso, essa comunidade busca se aproximar das práticas e objetivos propostos pela disciplina Humanidades Digitais.

**Palavras-chave:** Modelo de dados. Metadados. Representação da informação. *Linked Open Data*. Patrimônio cultural.

## ABSTRACT

**Introduction:** The emergence of the Internet and the Web platform raise new possibilities and challenges of management and access to knowledge, stored in digital environments in the form of data. In this context, the tools for description promoted by the Linked Open Data movement are highlighted, through Linked Data principles. The connection and availability of cultural heritage data, present in libraries, archives and museums, offer new possibilities for learning and dissemination of knowledge in Humanities and the vast cultural heritage distributed throughout the globe. Given this scenario, the question is: how do Linked Data principles contribute to the representation and access to cultural heritage collections?

**Purpose:** To analyze the Linked Open Data in cultural heritage institutions, in order to contribute to studies about representation and access to cultural heritage collections.

**Methodological procedures:** This is an exploratory theoretical-applied research, with a qualitative approach, carried out through a bibliographical and documentary survey about the research question. The scientific bibliography and relevant documents were analyzed in order to provide support for the representation and promotion of access to the cultural heritage collections within the scope of the Linked Open Data movement.

**Results:** The results achieved included 1) academic contributions regarding the benefits of using the tools of the Linked Open Data movement, highlighting issues of interoperability (technical and semantic) and data sustainability, 2) the survey of Linked Open Data initiatives in international cultural heritage institutions, 3) the presentation and discussion of data models adopted by the initiatives Europeana and DPLA and the data scheme of the DigitalNZ project and 4) presentation of the solutions developed by the cultural heritage community to use the metadata standards for the Linked Open Data description of cultural heritage collections. These results demonstrated that the cultural heritage community has sought to take advantage of the benefits offered by the Linked Open Data movement, especially with respect to the sustainability of the processes of description and access to the data of cultural heritage collections and, with that, this community seeks to approach the practices and goals proposed by the Digital Humanities discipline.

**Key-words:** Data model. Metadata. Cultural heritage. Information representation. Linked Open Data.

## LISTA DE ILUSTRAÇÕES

<b>Figura 1</b> - Arquitetura de ferramentas da proposta Web Semântica	30
<b>Figura 2</b> - Diagrama de princípios, movimentos e iniciativas relativas ao Linked Data	43
<b>Figura 3</b> - Ranking “5 estrelas” para publicação e ligação de dados	44
<b>Figura 4</b> - Visualização de datasets do projeto LOD-Cloud	46
<b>Figura 5</b> - Exemplos de 303 URI e hash URI	48
<b>Figura 6</b> - Fragmento de um documento XML	50
<b>Figura 7</b> - Dataset Europeana no projeto LOD-Cloud	62
<b>Figura 8</b> - Classes principais do modelo EDM	70
<b>Figura 9</b> - Exemplo de uso das classes contextuais do EDM	71
<b>Figura 10</b> - Esquema de conversão de metadados para o modelo EDM	72
<b>Figura 11</b> - Classes principais e contextuais do modelo DPLA MAP	73



## LISTA DE QUADROS

<b>Quadro 1</b> - Lista de fontes e termos para o levantamento bibliográfico	18
<b>Quadro 2</b> - Tipologia dos metadados	34
<b>Quadro 3</b> - Atributos e características dos metadados	35
<b>Quadro 4</b> - Padrões de metadados para patrimônio cultural	36
<b>Quadro 5</b> - Serializações RDF e suas funções	53
<b>Quadro 6</b> - Comparação entre os projetos Europeana, DPLA e DigitalNZ	77
<b>Quadro 7</b> - Desenvolvimento dos padrões de estrutura, conteúdo e de formato e intercâmbio de dados de coleções de patrimônio cultural	80
<b>Quadro 8</b> - Soluções para uso dos padrões de valores de dados na descrição Linked Open Data de coleções de patrimônio cultural	83

## SUMÁRIO

<b>1 INTRODUÇÃO</b>	11
1.1 Objetivos	13
1.2 Justificativa	13
1.3 Delimitação epistemológica da pesquisa	15
1.4 Procedimentos metodológicos	17
1.5 Estrutura do trabalho	19
<b>2 HUMANIDADES DIGITAIS E O PATRIMÔNIO CULTURAL</b>	21
<b>3 REPRESENTAÇÃO DE RECURSOS NO AMBIENTE WEB</b>	29
3.1 Dados e Metadados	32
3.2 Os princípios <i>Linked Data</i> e o movimento <i>Linked Open Data</i>	41
3.3 <i>Linked Open Data</i> e a descrição do patrimônio cultural	47
<b>4 SUSTENTABILIDADE DE DADOS E INICIATIVAS LINKED OPEN DATA EM INSTITUIÇÕES DE PATRIMÔNIO CULTURAL</b>	55
<b>5 MODELOS DE DADOS E PADRÕES DE METADADOS PARA DESCRIÇÃO LINKED OPEN DATA DE COLEÇÕES DE PATRIMÔNIO CULTURAL</b>	67
5.1 Modelos de dados	67
5.2 Padrões de metadados para descrição <i>Linked Open Data</i> de patrimônio cultural	79
<b>6 CONSIDERAÇÕES FINAIS</b>	86
<b>REFERÊNCIAS</b>	90

## 1 INTRODUÇÃO

O advento da *Internet* e da plataforma *Web*, aliados ao surgimento e desenvolvimento de outras tecnologias de informação e comunicação (TICs), protagonizaram o fenômeno do aumento exponencial da produção e disseminação de informações e conhecimentos provenientes de diversos domínios, o que por sua vez suscita novas possibilidades e desafios de gerenciamento e acesso a esses conhecimentos, armazenados nos ambientes digitais na forma de dados.

Nesse contexto, destaca-se o movimento *Linked Open Data*, ação do *World Wide Web Consortium* (W3C) que consiste na aplicação conjunta dos princípios *Linked Data* e *Open Data* (ARAKAKI, 2016). Estes princípios tratam da criação de relacionamentos semânticos entre dados por meio de *links* (BERNERS-LEE, 2006). O movimento *Linked Open Data*, então, traz diretrizes para o modo como dados devem ser disponibilizados para uso e indica ferramentas desenvolvidas para esse fim. Nesse sentido, a sociedade civil e as organizações têm acesso aos dados, e, também são capazes de manipulá-los e relacioná-los entre si, gerando novos conhecimentos.

As influências e benefícios da utilização dessas, e de outras ferramentas computacionais, também abarcam as Ciências Humanas e Sociais, sob a instituição de uma disciplina denominada Humanidades Digitais. Com características essencialmente práticas e interdisciplinares, esse novo campo de estudo e de trabalho trata da inserção de tecnologias nos fazeres dessas ciências, ao mesmo tempo em que busca a utilização dos conhecimentos em Humanidades na construção de novas ferramentas. Nesse viés, o Grupo de Pesquisa Humanidades Digitais (2010) da Universidade de São Paulo afirma que humanistas e tecnólogos devem efetivamente trabalhar juntos, pois nenhum deles, separadamente, conseguiria construir uma ferramenta para uso das Humanidades que atendesse a todos os requisitos técnicos e funcionais necessários.

A interdisciplinaridade e os aspectos práticos dessa disciplina se refletem, ainda, nos diversos eventos e grupos de trabalho que se formam para estudo e discussão de temáticas pertinentes ao campo. No Brasil, destaca-se o I Congresso Internacional em Humanidades Digitais, realizado em 2018 no Rio de Janeiro, que reuniu profissionais e cientistas de diversas áreas das Ciências Humanas e Sociais, inclusive da Ciência da Informação. Isso porque, dentre as temáticas das Humanidades Digitais encontram-se questões de divulgação, acesso e uso de informações culturais.

Na Ciência da Informação, tais questionamentos direcionam-se aos campos da organização e de representação da informação, historicamente alinhados com a aplicação das tecnologias da informação (ARAÚJO, 2014) e com especial atenção às questões de interoperabilidade entre sistemas de representação e recuperação de informações, utilizados por bibliotecas, arquivos e museus (MARCONDES, 2012, 2016) e de ampliação do acesso por parte dos usuários. Nesse contexto, tais instituições são também conhecidas como instituições *LAM* (da expressão em inglês "*libraries, archives and museums*").

Ao mesmo tempo, a literatura científica em Ciência da Informação considera as coleções dessas instituições como fontes de informação possíveis de se configurarem como patrimônio cultural (ARARIPE, 2004; SOUZA; CRIPPA, 2010). Assim, as relações das Humanidades Digitais com a Ciência da Informação se reforçam pelo aspecto cultural da informação e dos objetos que a registram e ampliam as questões patrimoniais para além do campo museológico.

Nessa perspectiva, coleções de instituições *LAM* devem ter seu acesso garantido e preservado a longo prazo. Para isso, os processos de representação devem buscar promover o acesso persistente aos dados gerados pelas descrições de coleções culturais. Dessa forma, é possível que o acesso persistente permita a interoperabilidade de dados, inclusive quando se trata de dados gerados pelas representações do patrimônio cultural.

Dentre as tecnologias que permitem a interoperabilidade de dados, as que se relacionam ao movimento *Linked Open Data* ganham destaque, uma vez que permitem o acesso, uso e reuso coerente de dados alocados em diferentes ambientes. Em se tratando de dados de patrimônio cultural, a conexão e disponibilização desses tipos de dados oferecem novas possibilidades de aprendizado e de divulgação dos conhecimentos em Humanidades e do vasto patrimônio cultural distribuído pelo globo.

Diante desse cenário, questiona-se: como os princípios *Linked Data* contribuem para a representação e o acesso às coleções de patrimônio cultural? Esta pesquisa, então, aborda essas temáticas, com vistas à inserção de instituições de patrimônio cultural nacionais no movimento *Linked Open Data*.

## 1.1 Objetivos

O objetivo geral da pesquisa é analisar o *Linked Open Data* em instituições de patrimônio cultural, no intuito de contribuir para os estudos da comunidade *LAM* acerca da representação e acesso às coleções de patrimônio cultural.

Como objetivos específicos, busca-se:

- Caracterizar os relacionamentos do patrimônio cultural com a Ciência da Informação e as Humanidades Digitais;
- Conceituar os princípios *Linked Data* e *Open Data*, o movimento *Linked Open Data* e as ferramentas de descrição de recursos no ambiente *Web*;
- Pontuar as iniciativas *Linked Open Data* em instituições de patrimônio cultural identificadas na literatura;
- Explorar as iniciativas Europeia, DPLA e *DigitalNZ*, considerando o seu escopo de atuação, seus modelos de dados e as tecnologias utilizadas;
- Descrever as soluções desenvolvidas pela comunidade de patrimônio cultural para a descrição *Linked Open Data* de suas coleções.

## 1.2 Justificativa

O movimento *Linked Open Data* é uma demanda global com implicações nos campos político, científico e também cultural. Nesse sentido, justifica-se esse trabalho pela sua contribuição às instituições de patrimônio cultural que estão identificando a necessidade de compreender tais implicações. Dentre elas, Rasmussen-Pennington (2016) destaca a maior visibilidade aos conteúdos de seus recursos representados na *Web*, que podem ser compreendidos também pelas máquinas por meio do uso dos princípios *Linked Data* e da proposta *Web Semântica* como um todo.

Ao mesmo tempo, os princípios *Linked data* oferecem às instituições culturais ferramentas e formas semânticas e interoperáveis de gerenciamento de dados, aumentando, assim, a quantidade e a qualidade de suas informações (BYRNE; GODDARD, 2010). Desse modo, o número de iniciativas *Linked Open Data* relativas ao patrimônio cultural, dentre elas

os projetos Europeia, *DigitalNZ* e *Digital Public Library of America* (DPLA), demonstra a dimensão que o conceito vem tomando em diferentes países.

Nessa perspectiva, a literatura científica internacional apresenta contribuições para a temática de inserção de dados de patrimônio cultural no escopo do movimento *Linked Open Data*, em que se destacam os trabalhos de Hyvönen (2012) e Van Hooland e Verborgh (2014). No âmbito nacional, Santos Neto (*et al.*, 2013), Arakaki (2016), Costa (2017), Romanetto (2017) e Assumpção (2018) apresentam contribuições variadas sobre os princípios *Linked Data* e o movimento *Linked Open Data*.

Pela literatura, também se observa que a integração dos três contextos, relativos às bibliotecas, arquivos e museus, permite maior valorização desses ambientes conjuntos, bem como a exploração da temática pelos trabalhos de Pós-Graduação Nacionais em Ciência da Informação. Desse modo, pesquisadores e grupos de pesquisa nessa área ainda possuem inúmeras formas de exploração do tema em consonância com o escopo maior dos estudos das Humanidades Digitais e do universo da ligação de dados.

A inclusão do Brasil nesse contexto só tem a contribuir para o mapeamento e gestão do patrimônio cultural nacional. Isso porque o país parece não acompanhar o cenário internacional. O Ministério da Cultura (MinC), bem como o Ministério da Justiça e Segurança Pública (MJSP) – responsável pelo Arquivo Nacional – anunciaram seus respectivos planos de abertura de dados, mas ambos não fazem menções significativas a respeito do patrimônio cultural dos quais são responsáveis (BRASIL, 2017a, 2017c). O Instituto Brasileiro de Museus (IBRAM), por outro lado, lidera o projeto Tainacan, que tem como objetivo o desenvolvimento de uma ferramenta que promova “a gestão de acervos culturais de arquivos, museus, bibliotecas e cinematecas, trazendo ainda uma camada voltada para a colaboração de usuários” (IBRAM, [201-], *online*). Entende-se, portanto, que as instituições brasileiras, sejam elas públicas ou particulares, científicas, políticas, educacionais e culturais ainda estão amadurecendo suas iniciativas de inserção ao movimento *Linked Open Data*.

Ao mesmo tempo, o advento da Lei de Acesso a Informação (LAI) foi impulsionado pelo processo de publicação dos dados, em consonância aos princípios *Open Data*, principalmente no meio governamental. No entanto, a lei mostra-se naturalmente limitada frente às propostas do movimento *Linked Open Data*. Para Santarem Segundo (2015) a estruturação dos dados de acordo com esse movimento permite uma aplicação mais efetiva da LAI, em âmbito nacional, uma vez que permite que os dados disponibilizados sejam processados por máquina.

Para a sociedade, a disponibilização de dados ligados de coleções de patrimônio cultural favorece o enriquecimento social e cultural dos cidadãos, pois facilita diferentes modos de acesso, uso e reuso dos dados disponibilizados.

Finalmente, o interesse pessoal da autora em relação à cultura, suas manifestações e objetos, bem como o interesse acadêmico em relação à representação desse tipo de coleção são motivadores para a elaboração e execução desta pesquisa dentro do Programa de Pós-Graduação em Ciência da Informação da Universidade Federal de São Carlos (UFSCar).

### **1.3 Delimitação epistemológica da pesquisa**

Uma vez que esta pesquisa se fundamenta em temas emergentes, quais sejam, o movimento *Linked Open Data* e as Humanidades Digitais, ambos sob o escopo da Ciência da Informação, e que esta possui características de surgimento, evolução e objetos de estudo, observa-se a necessidade de uma delimitação epistemológica deste trabalho.

Dentre as considerações sobre o que se configura a Ciência da Informação, qual é o seu objeto de estudo e a sua relevância para a sociedade, esta pesquisa adota como parâmetro as definições e discussões apresentadas por Araújo (2014), Buckland (1991), Capurro (2003), Capurro e Hjørland (2007), Castells (2007), Freire (2006), Saracevic (1996) e Silva e Freire (2012).

Buckland (1991) define a informação por meio de sua divisão em três aspectos: informação como coisa - objetos que são informativos -, informação como processo - ato de informar/comunicar - e informação como conhecimento - aquilo que é percebido ou absorvido no processo de comunicação. Percebe-se que as questões relativas ao suporte informacional são evidenciadas, assim como se consideram os demais aspectos relativos à informação e aos fluxos informacionais. Cada um desses aspectos configura diferentes escopos teóricos e práticos dentro da Ciência da Informação.

Para Saracevic (1996) a Ciência da Informação é uma ciência pautada na interdisciplinaridade, nas tecnologias de informação e na sociedade da informação, estudando o modo de criação, uso e comunicação da informação a partir de ligações com a Ciência da Computação, a Ciência Cognitiva, a Biblioteconomia e a Comunicação. De modo mais abrangente, o armazenamento e recuperação de informações (FREIRE, 2006), a organização e disseminação de informações, a organização do conhecimento (SILVA; FREIRE, 2012) e a

criação e expansão das TICs (CAPURRO; HJORLAND, 2007) são questões que permeiam a atuação dos cientistas da informação.

Araújo (2014), Capurro (2003) e Capurro e Hjørland (2007) discutem, então, os paradigmas dessa ciência, de acordo com os conceitos de informação apresentados por Buckland (1991). Capurro (2003) e Capurro e Hjørland (2007) utilizam-se diretamente do termo paradigma para descrever os conceitos de informação dentro da Ciência da Informação: paradigma físico, paradigma cognitivo e paradigma social. Posteriormente, Araújo (2014, p.70) utiliza-se dessas definições para resumir que

[...] o primeiro conceito de informação na Ciência da Informação é mais restrito e está vinculado à sua dimensão material, física, sendo o fenômeno estudado a partir de uma perspectiva quantitativa e positivista. Nos anos seguintes, tomou corpo um conceito um pouco mais amplo voltado para a dimensão cognitiva, sendo informação algo associado à interação entre dados (aquilo que existe materialmente) e conhecimento (aquilo que está na mente dos sujeitos), e seu estudo relacionado à identificação de significados, interpretações. Por fim as tendências contemporâneas implicam um grau maior de complexidade e abstração, com a inserção da informação no escopo da ação humana e no âmbito de contextos socioculturais concretos.

Embora a literatura geralmente trate dos paradigmas apontados como revoluções científicas lineares e cronológicas, Capurro (2003) esclarece que as ideias do denominado paradigma social já estavam embutidas nas práticas biblioteconômicas e documentárias tradicionais e anteriores ao surgimento da Ciência da Informação. Ao mesmo tempo, ao trazer os aspectos de informação como coisa, como processo e como conhecimento, Buckland (1991) está na verdade se utilizando dos três conceitos paradigmáticos para determinar o escopo e a atuação da Ciência da Informação nas últimas décadas. Isso porque, o advento da Internet, em um curto espaço de tempo, trouxe soluções e ao mesmo tempo potencializou os problemas de armazenamento, representação, acesso, uso e comunicação de informações, incorporados ao ambiente *Web*.

A proposta *Web Semântica* e seus desdobramentos, em especial o movimento *Linked Open Data*, trouxeram implicações para o campo da organização e representação de informações em ambientes digitais, área de estudo desta pesquisa. Trata-se aqui, então, do aspecto de informação como coisa, uma vez que os sistemas de informação lidam apenas com esse tipo de informação (BUCKLAND, 1991). Ao mesmo tempo, ao se considerar que a Ciência da Informação está intimamente relacionada à sociedade em rede, como apontado por Castells (2007), e que as novas tecnologias trazem implicações cognitivas (interação com as máquinas) e de relações sociais (redes sociais, por exemplo), é razoável inferir que,



atualmente, os três grandes conceitos ou paradigmas de informação coexistem no mesmo espaço-tempo e determinam a atuação da Ciência da Informação nestas três frentes.

Considera-se, então, que esta pesquisa se insere no paradigma físico da Ciência da Informação, ao mesmo tempo em que possui diálogos adjacentes com questões cognitivas e sociais da informação, expressos nas relações com o campo das Humanidades Digitais.

#### **1.4 Procedimentos metodológicos**

A pesquisa possui natureza teórico-aplicada, uma vez que tem como finalidade “[...] gerar conhecimentos para aplicação prática, dirigidos à solução de problemas específicos” (GERHARDT; SILVEIRA, 2009, p.35). Neste caso, busca-se analisar a perspectiva *Linked Open Data* em instituições de patrimônio cultural para sua inserção em outras instituições da comunidade *LAM*.

Devido aos seus objetivos, é uma pesquisa exploratória, que para Gil (2002, p.41) “[...] tem como objetivo proporcionar maior familiaridade com o problema, com vistas a torná-lo mais explícito ou a construir hipóteses” e de abordagem qualitativa. Assim, procura-se identificar características do objeto, o patrimônio cultural, ao mesmo tempo em que se procura compreender suas particularidades em relação ao problema de pesquisa: questões de representação e acesso frente às possibilidades oferecidas pelo movimento *Linked Open Data*.

Quanto aos seus procedimentos, envolveu levantamento bibliográfico e documental acerca de sua temática. As buscas por literatura científica em bases de dados especializadas, nacionais e internacionais, foram realizadas de acordo com o quadro 1, por meio do uso dos termos “*linked data*”, “*linked open data*”, “*digital humanities*” e “*cultural heritage*”, bem como seus correspondentes em português. Eventualmente, utilizou-se o termo “*Semantic Web*” ou “*Web Semântica*” para a recuperação de trabalhos que abordam as ferramentas pertinentes ou que forneçam uma contextualização do tema desta pesquisa. Ressalta-se também a utilização de operadores de refinamento e de filtros disponíveis em cada das fontes elencadas de acordo com as especificidades dos assuntos em relação ao escopo desse trabalho.

**Quadro 1** - Lista de fontes e termos para o levantamento bibliográfico

Fontes		Termos	Data de publicação
Base de Dados Referenciais de Artigos de Periódicos em Ciência da Informação (BRAPCI)		<i>Linked Open Data</i> <i>Linked Data</i> Patrimônio cultural Humanidades Digitais	Janeiro de 2000 a Março de 2018
Repositório Institucional da UNESP		<i>Linked Data</i> <i>Linked Open Data</i>	
Encontro Nacional de Pesquisa em Ciência da Informação (ENANCIB)		<i>Linked Data</i> <i>Linked Open Data</i> Patrimônio cultural	
Biblioteca Digital Brasileira de Teses e Dissertações (BDTD)		<i>Linked Data</i> <i>Linked Open Data</i> Patrimônio cultural	
<i>Scientific Electronic Library Online (SciELO)</i>		<i>Linked Open Data</i> <i>Linked Data</i> Patrimônio cultural Humanidades Digitais	
Portal de Periódicos da CAPES	<i>Library and Information Science Abstracts (LISA)</i>	<i>Linked Open Data</i> <i>Linked Data</i> <i>Cultural heritage</i> <i>Digital Humanities</i>	
	<i>Information Science &amp; Technology Abstracts (ISTA)</i>	<i>Linked Open Data</i> <i>Linked Data</i> <i>Cultural heritage</i> <i>Digital Humanities</i>	
	<i>Library, Information Science &amp; Technology Abstracts</i>	<i>Linked Open Data</i> <i>Linked Data</i> <i>Cultural heritage</i> <i>Digital Humanities</i>	
<i>Web of Science</i>		<i>Linked Open Data</i> <i>Linked Data</i> <i>Cultural heritage</i>	

Fonte: elaborado pela autora

Embora esse levantamento tivesse como objetivo fornecer subsídios para o referencial teórico da pesquisa, o levantamento documental também contribuiu para esse fim. Assim, os mesmos termos foram utilizados para o levantamento documental pertinente ao escopo do trabalho, destacando-se a documentação referente às iniciativas *Linked Open Data* de instituições de patrimônio cultural selecionadas, às ferramentas abordadas ao longo do

trabalho e às soluções desenvolvidas pela comunidade de patrimônio cultural para uso dessas ferramentas em conjunto aos seus instrumentos de representação.

As iniciativas *Linked Open Data* apresentadas foram selecionadas por meio da literatura científica, tendo como critério básico de inclusão nos resultados da pesquisa a integração dos contextos bibliográfico, arquivístico e museológico; ou seja, foram selecionadas as iniciativas que contemplam os três tipos de coleções. Ao mesmo tempo, os projetos que fornecem acesso aos seus dados via *Application Programming Interface* (APIs) foram selecionados para discussão e análise comparativa de aspectos específicos de sua modelagem de dados. Os projetos Europeana, *Digital Publica Library of America* (DPLA) e *DigitalNZ* atenderam à esse critério.

A bibliografia científica e os documentos pertinentes foram analisados de modo a fornecer subsídios para a inserção de instituições *LAM* no movimento *Linked Open Data* por meio do uso das ferramentas desse movimento para a representação e promoção do acesso às suas coleções de patrimônio cultural. Os resultados trataram respectivamente:

- dos benefícios percebidos no uso dessas ferramentas;
- das contribuições acadêmicas à temática identificadas na literatura;
- das iniciativas *Linked Open Data* em instituições de patrimônio cultural identificadas na literatura;
- dos modelos de dados adotados pelas iniciativas Europeana e DPLA e do esquema de dados do projeto *DigitalNZ* e
- das soluções desenvolvidas pela comunidade de patrimônio cultural para descrição *Linked Open Data* de suas coleções.

Estes resultados foram apresentados na forma de relatório textual, figuras e quadros de comparação, segundo a estrutura do trabalho, apresentada a seguir.

## **1.5 Estrutura do trabalho**

As temáticas apresentadas e discutidas ao longo do trabalho estão distribuídas em capítulos, conforme especificado abaixo:

- O capítulo inicial delimita o objeto, o tema e a questão de pesquisa, apresenta seus objetivos, justificativa, delimitação epistemológica e procedimentos metodológicos.
- **Capítulo 2** – Apresenta considerações sobre a disciplina denominada Humanidades Digitais e suas relações com a Ciência da Informação, bem como questões relativas ao patrimônio cultural.
- **Capítulo 3** – Discute questões sobre a representação de informações em ambientes digitais. Apresenta os conceitos de dados e metadados, elencando os padrões de metadados desenvolvidos e utilizados pela comunidade de patrimônio cultural. Apresenta e discute os conceitos, características e diferenças entre os princípios *Linked Data* e *Open Data* e o movimento *Linked Open Data*. Apresenta ferramentas e tecnologias utilizadas por esse movimento.
- **Capítulo 4** – Apresenta o levantamento de iniciativas *Linked Open Data* em instituições de patrimônio cultural e o cenário brasileiro de ligação e disponibilização de dados de patrimônio cultural.
- **Capítulo 5** – Apresenta e discute os modelos de dados desenvolvidos pelas iniciativas Europeia, DPLA e *DigitalNZ* e as soluções desenvolvidas para uso dos padrões de metadados na descrição *Linked Open Data* de coleções de patrimônio cultural.
- **Capítulo 6** – Considerações finais do trabalho e indicações de estudos futuros.

## 2 HUMANIDADES DIGITAIS E O PATRIMÔNIO CULTURAL

O protagonismo e as possibilidades cada vez mais sofisticadas para o desenvolvimento e uso das TICs na sociedade moderna têm exercido influências na área de Ciência da Informação, bem como em outras áreas interdisciplinares.

Nesse viés, as Ciências Sociais Aplicadas e as Ciências Humanas estão imersas na idealização de uma nova conjunção de estudo e de trabalho, denominada Humanidades Digitais. Ainda em franco desenvolvimento, o termo designa genericamente um movimento coletivo, inclusive de centros de informação e memória - como arquivos, bibliotecas, museus e galerias de arte - para inserção de diferentes tecnologias na sua atuação diária.

Por ser um campo emergente, a literatura registra diferentes questionamentos e discussões sobre o que se configura, ou não, como escopo das Humanidades Digitais e se esse novo campo possui aspectos científicos: Almeida e Damian (2015) consideram as Humanidades Digitais como um campo interdisciplinar que possui questões teóricas e filosóficas próprias – o que lhe atribuiria um *status* de ciência - enquanto Koltay (2016) afirma que é um campo voltado prioritariamente à aplicação de métodos - o que lhe confere aspectos práticos.

Independentemente dos questionamentos pelos quais essa disciplina se submete, a literatura é unânime em considerá-la como interdisciplinar ou transdisciplinar, uma vez que não se sustenta sem o estabelecimento de pontes com outras disciplinas, ou outras ciências, como por exemplo, a Ciência da Informação.

Koltay (2016) esclarece que a fundamentação teórica das Humanidades Digitais foi estabelecida pelas Humanidades 'tradicionais'. Esta afirmação encontra apoio na definição registrada no manifesto publicado pelo *The Humanities and Technology Camp* (THATCamp), no qual define-se que as “Humanidades Digitais designam uma ‘transdisciplina’, isto é, uma portadora dos métodos, dos dispositivos e das perspectivas heurísticas ligadas ao digital no domínio das Ciências Humanas e Sociais” (DACOS, 2011, *online*, tradução nossa<sup>1</sup>).

O THATCamp denomina uma conferência informal realizada em várias partes do mundo, reunindo cientistas e técnicos das ciências Sociais e Humanas e da área de tecnologia (DACOS, 2011) para estudo e discussão de temas das Humanidades Digitais. No Brasil, o I Congresso Internacional em Humanidades Digitais (HDRio) reuniu acadêmicos, cientistas e

---

<sup>1</sup> *The digital humanities designate a “transdiscipline”, embodying all the methods, systems and heuristic perspectives linked to the digital within the fields of humanities and the social sciences.*

tecnólogos das Artes, da Cultura e das Ciências Sociais, Humanas, Exatas e Computacionais para discussões acerca dessa disciplina no cenário nacional. Este evento afirma que as Humanidades Digitais,

[...] enquanto expressão agregadora de práticas, teorias e métodos que se desenvolvem desde a segunda metade do século XX, são consideradas, pelos muitos laboratórios e centros de pesquisa surgidos desde 2010 no mundo acadêmico, como uma “transdisciplina”, que incorpora os métodos, os dispositivos e as perspectivas heurísticas das ciências humanas e sociais, ao mesmo tempo em que mobiliza as ferramentas e abordagens singulares abertas pela tecnologia digital (CONGRESSO INTERNACIONAL EM HUMANIDADES DIGITAIS, 2018, *online*)

Percebe-se que tal afirmação converge com as definições do Dacos (2011) e Koltay (2016) no que diz respeito à absorção de teorias das ciências Sociais e Humanas, o que confere às Humanidades Digitais caráter interdisciplinar e transdisciplinar. Os dois eventos são ainda, exemplos de grupos de estudantes, profissionais ou cientistas que se unem com a finalidade de se discutir as temáticas pertinentes ao campo e reforçam os aspectos de interdisciplinaridade e transdisciplinaridade registrados na literatura.

Porém, autores como Doueihi (2010), Liu (2012), Almeida e Damian (2015), Koltay (2016) e Revez (2017) demonstram que o escopo das Humanidades Digitais está além da simples absorção de tecnologias pelas Ciências Humanas e Sociais.

Douihi (2010) considera que essa disciplina trata de todo um ‘entorno digital’, sendo este compreendido por tecnologias, instrumentos digitais, usos, práticas e marcos jurídicos que governam as relações entre todos esses elementos. Nesse sentido, pode-se dizer que esse ‘entorno digital’ se trata de um sistema tecnológico complexo, cujos atores e componentes não podem ser abordados isoladamente.

Fundamentando-se em Doueihi (2010), Almeida e Damian (2015) afirmam que as Humanidades Digitais envolvem questões relativas à apropriação da informação e aos aspectos de empoderamento (*empowerment*) gerados por ações e políticas culturais, justificando que estas se transformam posteriormente em iniciativas e experimentos de criação e construção de competências e habilidades relacionadas aos usos possíveis das tecnologias de informação. Nesse caso, a abordagem volta-se para os aspectos sociais da aplicação das tecnologias dentro desse sistema.

Ao mesmo tempo, Revez (2017) reitera em sua pesquisa a diferença entre avaliar o uso das tecnologias e perceber a apropriação de um discurso de viés tecnológico que altera procedimentos investigativos e, como consequência, a interpretação dos fenômenos humanos

e sociais. Pode-se inferir, então, que Revez (2017) categoriza as pesquisas em Humanidades Digitais de acordo com aspectos de avaliação quantitativos e qualitativos do uso das tecnologias pelas Ciências Humanas e Sociais.

Assim, entende-se que as Humanidades Digitais tratam também dos usos das tecnologias como novos meios de produção de conhecimentos, especialmente em humanidades. Koltay (2016) considera que as Humanidades Digitais buscam compreender os efeitos da infraestrutura digital no sistema de produção de conhecimento das humanidades e os modos pelos quais esse conhecimento é traduzido novamente em informação, por meio das técnicas computacionais, para sua disseminação.

No mesmo sentido, Liu (2012) e Koltay (2016) ressaltam que essa disciplina é uma evolução de outra precedente, denominada como Humanidades Computacionais. Isso por que relaciona a linguística computacional, os métodos tradicionais de análise, codificação de textos e de construção de ferramentas aos estudos das novas mídias, dos processos de comunicação e seus impactos na sociedade, reforçando as relações conceituais entre tecnologia, mídia e informação (LIU, 2012). Ressalta-se, então, que esses elementos são inseparáveis e que relacionamentos entre eles são fundamentais para o entendimento dos fluxos da informação na sociedade atual.

Ainda para Liu (2012), o alinhamento entre os estudos ditos tradicionais com os contemporâneos expande o denominado paradigma interpretativo, pois atualmente computadores e humanos compartilham responsabilidades cognitivas. Este paradigma relaciona-se diretamente com os objetivos da proposta *Web Semântica* e as ferramentas desenvolvidas para tal finalidade.

Spiro (2012, p.4) pontua as considerações sobre a temática das Humanidades Digitais em objetivos, que são: a) promover amplo acesso à informação cultural; b) permitir a manipulação de dados; c) transformar a comunicação acadêmica; d) melhorar os processos de ensino-aprendizagem e e) ter impacto público. Resumidamente, esses objetivos convergem para a democratização da criação, do acesso, do uso e reuso do conhecimento gerado por essas ciências, utilizando as tecnologias em todas as etapas do sistema de produção de conhecimento vigente para as Humanidades.

Em consonância com a tríade de tecnologia, mídia e informação de Liu (2012), esses objetivos se traduzem ainda em atividades e conceitos diretamente relacionados à Ciência da Informação praticada nas últimas décadas e voltada para a Sociedade da Informação, como apontado por Saracevic (1995). Ao mesmo tempo, enquanto Ciência Social Aplicada, a

Ciência da Informação contribui aos estudos ainda em consolidação das Humanidades Digitais.

Koltay (2016) traça uma similaridade entre as duas vertentes ao afirmar que as Humanidades Digitais têm como escopo a investigação, a análise, a síntese e a apresentação da informação registrada em formato digital. Percebe-se a semelhança entre essa afirmação e a definição geral sobre o escopo da Ciência da Informação, bem como a referência ao aspecto de informação como coisa, definida e explicada por Buckland (1991) como sendo a única manipulável por sistemas de informação.

Ao mesmo tempo, Porsdam (2013) elenca a dimensão cultural da computação como ponto de exploração para as duas disciplinas, ultrapassando o entendimento tradicional da cultura por meio da tecnologia. Assim, Koltay (2016) e Porsdam (2013) tratam de aspectos relacionados aos paradigmas físico e social da Ciência da Informação em consonância com questionamentos atuais das Humanidades Digitais.

No nível institucional, Vandegrift (2012) resume que essa disciplina e os centros de informação têm em comum as questões de acessibilidade da informação, ao passo que Marcondes (2016) atenta para a possibilidade real de alcance da almejada interoperabilidade entre sistemas de instituições *LAM* a partir do uso específico das ferramentas desenvolvidas para a proposta *Web Semântica*, permitindo o estabelecimento de relações culturalmente significativas entre suas coleções e, portanto, o reconhecimento do valor dessas relações.

Nesta afirmativa, Marcondes (2016) indica o salto qualitativo que essas ferramentas, ou tecnologias, oferecem aos catálogos de instituições *LAM*, que têm sido sistematicamente transformados pelas tecnologias e métodos propostos pela Ciência da Informação e áreas correlatas. Tais transformações naturalmente se refletem em outros processos dessas instituições e, de forma mais ampla, nos processos que envolvem essa ciência como um todo e que atendem aos objetivos propostos por Spiro (2012) em relação às Humanidades Digitais.

Contudo, o objetivo primeiro de promover o acesso à informação cultural contém particularidades até então pouco exploradas pela Ciência da Informação, uma vez que envolve questões relativas ao patrimônio cultural, tradicionalmente reservadas apenas aos museus (DODEBEI, 2006). Assim, faz-se necessário compreender o conceito de patrimônio e os processos que envolvem o acesso a esse tipo de objeto informacional.

Autores como Maroevic (1998), Souza e Crippa, (2010) e Dahlström, Hansson e Kjellman (2012) consideram que tanto os acervos museológicos, quanto os bibliográficos e os arquivísticos são passíveis de configurarem como patrimônio cultural: arquivos, assim como museus, lidam com objetos únicos, enquanto que as bibliotecas têm como objetivo reunir o



conhecimento humano (DAHLSTRÖM; HANSSON; KJELLMAN, 2012). Desse modo, as questões sobre o patrimônio ultrapassam o domínio dos museus e passam a fazer parte também dos âmbitos arquivístico e bibliotecário.

Dodebei (2006) esclarece que esta concepção é recente, devido ao surgimento e a multiplicação de centros de memória e de cultura, onde diferentes tipos de bens são absorvidos por essas instituições e passam ser considerados patrimônios. Conseqüentemente, são fontes de informação (ARARIPE, 2004; SOUZA; CRIPPA, 2010).

Gomes (2016) esclarece que o conceito primordial de patrimônio está relacionado ao conceito de propriedade, ou posse – motivo de debates e tensões entre o poder público e o privado, a sociedade e o indivíduo “dono” do bem - e que uma preocupação social, institucionalizada, em relação ao patrimônio, sua preservação, conservação e restauração, iniciou-se no período da Revolução Francesa, no século XVIII, voltada aos grandes monumentos.

A nível global, essa preocupação culminou na “Convenção sobre a proteção do Patrimônio Mundial Cultural e Natural” da Organização das Nações Unidas para a Educação, a Ciência e a Cultura (UNESCO), realizada em Paris no ano de 1972 (CAVALCANTE, 2007; SOUZA; CRIPPA, 2010) e ratificada, em 2002, pela “Declaração de Budapeste” (COMITÉ DO PATRIMÓNIO MUNDIAL, 2002). Esta convenção define patrimônio cultural como monumentos, conjuntos e sítios de valor universal excepcional do ponto de vista da história, arte, ciência e/ou estética, além do ponto de vista etnológico ou antropológico no caso dos sítios (INSTITUTO DO PATRIMÔNIO HISTÓRICO E ARTÍSTICO NACIONAL, 2008).

No Brasil, Souza e Crippa (2010) afirmam que os questionamentos sobre a identidade e o patrimônio cultural nacional começaram a ser abordadas pelos artistas vinculados ao Movimento Modernista Brasileiro, na década de 20 do século XX, sendo a arquitetura barroca considerada como alicerce do patrimônio efetivamente brasileiro. A criação do Serviço do Patrimônio Histórico e Artístico Nacional (SPHAN), no ano de 1937, com foco no patrimônio edificado (ARARIPE, 2004), parece ratificar a ideia. O mesmo ainda se reflete no Decreto-Lei nº25, do mesmo ano, que normaliza as questões sobre o patrimônio histórico e artístico nacional e traz a definição deste como “[...] o conjunto de bens móveis e imóveis existentes no país e cuja conservação seja de interesse público, quer por sua vinculação a fatos memoráveis da história do Brasil, quer por seu excepcional valor arqueológico ou etnográfico, bibliográfico ou artístico” (BRASIL, 1937).

O foco no patrimônio histórico parece esclarecer a ausência do conceito de patrimônio cultural no âmbito institucional até a Constituição de 1988, que traz a seguinte definição:

Constituem patrimônio cultural brasileiro os bens de natureza material e imaterial, tomados individualmente ou em conjunto, portadores de referência à identidade, à ação, à memória dos diferentes grupos formadores da sociedade brasileira, nos quais se incluem:

- I - as formas de expressão;
- II - os modos de criar, fazer e viver;
- III - as criações científicas, artísticas e tecnológicas;
- IV - as obras, objetos, documentos, edificações e demais espaços destinados às manifestações artístico-culturais;
- V - os conjuntos urbanos e sítios de valor histórico, paisagístico, artístico, arqueológico, paleontológico, ecológico e científico (BRASIL, 1988)

As definições da UNESCO e da Constituição Nacional evidenciam o caráter plural dos objetos que podem ser considerados patrimônio cultural e os diferentes dispositivos para a sua identificação e proteção. Maroevic (1998) lembra que o patrimônio do tipo material é dividido entre móvel e imóvel, enquanto Hyvönen (2012) remete à divisão entre tangível (objetos, artefatos, livros, etc.), intangível (linguagem, tradições, folclore, etc.) e natural (lugares, biodiversidade, etc.).

A Convenção da UNESCO de 1972 trata do patrimônio cultural dos tipos naturais e imóveis, enquanto que a atual legislação brasileira inclui também os bens móveis e intangíveis (imateriais). Percebe-se o histórico nacional e internacional das discussões a respeito do que se considera patrimônio, onde as preocupações iniciais se voltaram para os grandes monumentos construídos pelo homem, sendo posteriormente considerados os objetos móveis e finalmente, a própria cultura e sua dimensão imaterial.

Nessa perspectiva, Fonseca (2004) afirma que o Brasil é pioneiro em relação ao patrimônio cultural imaterial, pois foi um dos primeiros países a pensar no registro e proteção desse tipo de patrimônio, através de decreto presidencial do ano 2000<sup>2</sup>, anterior à “Convenção para a Salvaguarda do Patrimônio Cultural Imaterial”, proposta pela UNESCO em 2003 (SOUZA; CRIPPA, 2011). Esta convenção define o patrimônio imaterial como as “[...] práticas, representações, expressões, conhecimentos e técnicas - junto **com os instrumentos, objetos, artefatos e lugares culturais que lhes são associados** - que as comunidades, os grupos e, em alguns casos, os indivíduos reconhecem como parte integrante de seu patrimônio cultural” (UNESCO, 2006, não paginado, grifo nosso).

---

<sup>2</sup> Este decreto cria o Registro de Bens Culturais de Natureza Imaterial e o Programa Nacional do Patrimônio Imaterial (BRASIL, 2000).

A UNESCO, portanto, considera os objetos físicos que acompanham os fazeres e práticas culturais como parte integrante do patrimônio imaterial e, conseqüentemente, cultural. O mesmo pode-se dizer da Constituição Nacional, anterior à Convenção da UNESCO, e que no quarto item lista obras (domínio bibliográfico), objetos (domínio museológico) e documentos (domínio arquivístico) como possíveis constituintes do patrimônio cultural nacional. Assim, Hyvönen (2012, p. 01, grifo e tradução nossa<sup>3</sup>) define que patrimônio cultural se refere “[...] ao legado de **objetos físicos, ambientes, tradições e conhecimentos de uma sociedade**, que pertencem ao passado, que são mantidos e desenvolvidos no presente e que são preservados (conservados) para o benefício das gerações futuras”.

Nota-se, portanto, a relação intrínseca entre os aspectos materiais e imateriais que estão presentes no patrimônio cultural. As duas últimas definições afirmam o papel que a própria sociedade possui no reconhecimento daquilo que deve ser julgado como patrimônio ou não, pois segundo Araripe (2004), o patrimônio cultural diz respeito ao objeto (material) ou manifestação cultural (imaterial) que tem significado e sentido social.

O patrimônio imaterial possui sempre alguma materialidade, enquanto o material está revestido de uma imaterialidade que lhe dá sentido. Atributos materiais e imateriais estão presentes em todo e qualquer patrimônio (ARARIPE, 2004; SOUZA; CRIPPA, 2011), razão pela qual deve ser visto como um processo social e cultural de atribuição de sentido pelos usos sociais (SMITH, 2008).

Ao mesmo tempo, Maroevic (1998) afirma que é um processo dinâmico, pois “[...] na realidade é a harmonia entre matéria, forma e significado, que se manifesta no tempo, no espaço e na sociedade, que faz patrimônios culturais individuais autênticos” (p. 136, tradução nossa<sup>4</sup>). As variáveis como o tempo, espaço e sociedade foram profundamente transformadas pelo mundo digital, de modo que matéria, forma e significado assumem também outras conotações. Surge, então, uma nova categoria de patrimônio cultural, denominada patrimônio digital, que para Dodebei (2006, não paginado) refere-se aos

---

<sup>3</sup> “*Cultural Heritage (CH) refers to the legacy of physical objects, environment, traditions, and knowledge of a society that are inherited from the past, maintained and developed further in the present, and preserved (conserved) for the benefit of future generations*”.

<sup>4</sup> “*It is actually the concord of material, form and significance, which is manifested in time, space and society, that makes individual parts of cultural heritage authentic*”.

[...] bens culturais criados somente em ambiente virtual ou por bens duplicados na representação da *Web* e sobre materiais digitais que incluem textos, bases de dados, imagens estáticas e com movimento, áudios, gráficos, software, [sic] e páginas *Web*, entre uma ampla e crescente variedade de coleções que representam desde objetos pessoais a acervos tradicionais de instituições de memória.

O patrimônio digital, então, engloba não apenas a cópia digital de determinado objeto, mas todo um novo conjunto de dispositivos que surgiram com o desenvolvimento tecnológico, por exemplo, os *softwares* e as páginas *Web*.

Dodebei (2006) ainda elenca algumas questões relativas à migração da sociedade do ambiente analógico para o digital, em que se destacam a necessidade de salvaguarda e preservação desse tipo de patrimônio, no sentido de manter o seu significado e sua acessibilidade diante de um ambiente que se transforma a cada segundo. Para isso, a UNESCO elaborou o ‘Manifesto para a Preservação Digital’, cujas propostas incluem a elaboração de princípios de preservação e acesso ao patrimônio digital em colaboração direta com bibliotecas, arquivos e museus, bem como a promoção e o uso de normas e soluções tecnológicas desenvolvidas em conjunto à indústria de tecnologia da informação (UNESCO, 2003).

Ao convocar bibliotecas, arquivos e museus, a UNESCO atesta a importância que as instituições *LAM* possuem para a salvaguarda e acesso ao patrimônio digital, ao mesmo tempo em que evidencia a necessidade de adoção de ferramentas tecnológicas capazes de gerenciar o crescente número de informações, ou patrimônios, gerados no universo digital, sejam eles nativos digitais ou fruto de conversão do formato analógico para o digital.

A partir, então, da compreensão do patrimônio cultural como um conjunto total e coeso de objetos, práticas e conhecimentos e do reconhecimento da necessidade de representá-lo como tal, recorre-se às discussões relativas à representação de informações no ambiente digital, especificamente ao uso de metadados e de padrões de metadados, bem como de ferramentas de descrição utilizadas pelo movimento *Linked Open Data*, apresentadas e discutidas no próximo capítulo.

### 3 REPRESENTAÇÃO DE RECURSOS NO AMBIENTE WEB

A plataforma *Web*, assim como outros ambientes digitais, tem sido transformada pelas tecnologias desde sua criação e crescente popularização. No cenário dito tradicional da *Web*, destacam-se as linguagens de marcação como meios de apresentação das informações contidas nos documentos *Web*. Nesse contexto, Almeida (2002) define as linguagens de marcação como um conjunto de convenções utilizadas para a codificação de textos, que especificam as marcas a serem utilizadas e quais os seus significados.

A proposta *Web Semântica* tem como objetivo modificar esse cenário, por meio do uso de ferramentas padronizadas de representação semântica nas camadas mais básicas da plataforma, transformando a *Web* de documentos em uma *Web* de dados: “[...] para tornar a *Web* de dados uma realidade, é importante que se tenha uma grande quantidade de dados disponíveis na *Web* em formatos padronizados, recuperáveis e manuseáveis por ferramentas da *Web Semântica*” (WORLD WIDE WEB CONSORTIUM, 2015, não paginado, tradução nossa<sup>5</sup>).

Assim, as ferramentas desenvolvidas até o momento têm como objetivo realizar a descrição, e conseqüentemente, a representação semântica não somente dos documentos como um todo, mas dos dados contidos nos documentos e dos diversos tipos de *links* que existem entre eles (RAMALHO; VIDOTTI; FUJITA, 2009). Desse modo, os documentos são fragmentados em dados e estes são trabalhados de forma independente.

Tais ferramentas encontram-se distribuídas em diferentes camadas da *Web*. Na camada estrutural encontra-se a ferramenta *Uniform Resource Identifier* (URI), responsável pela identificação única de cada recurso na *Web*. É na camada sintática que se trabalha com as ferramentas responsáveis pela descrição dos recursos e de relacionamentos (*links*) entre eles: a *eXtensible markup language* (XML) e o modelo *Resource Description Framework* (RDF), considerado a linguagem geral de representação na *Web*.

A camada efetivamente semântica envolve o uso de vocabulários, gerais ou de domínio, e ontologias, ferramenta que Ramalho (2006, p. 97) define como “[...] um artefato tecnológico que descreve um modelo conceitual de um determinado domínio em uma linguagem lógica e formal, a partir da descrição de aspectos semânticos dos conteúdos informacionais”. As ferramentas desenvolvidas pelo consórcio W3C para essa camada são: *Ontology Web Language* (OWL) - um modelo geral para definir e instanciar ontologias na

---

<sup>5</sup> “[...] to make the *Web of Data* a reality, it is important to have the huge amount of data on the *Web* available in a standard format, reachable and manageable by *Semantic Web* tools”.

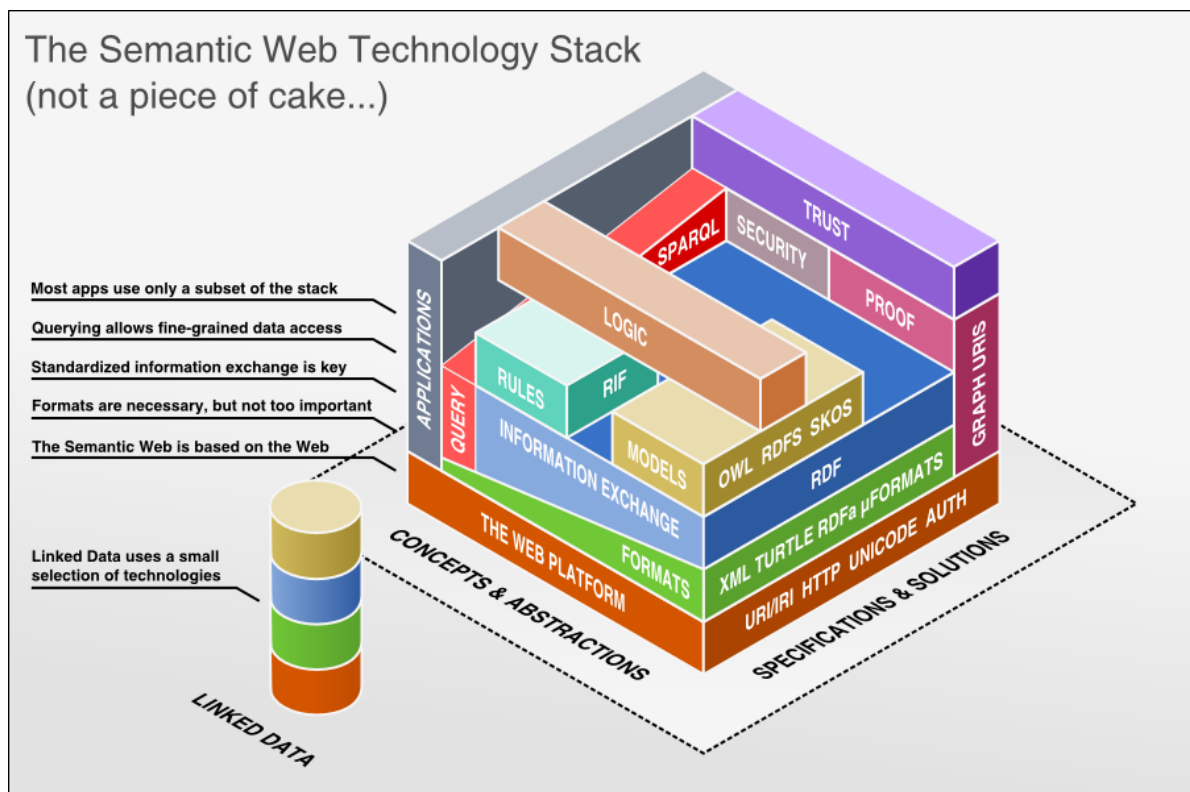
*Web* (MÉNDEZ RODRÍGUEZ; GREENBERG, 2012), *RDF Schema* (RDF-S) e *Simple Protocol and RDF Query Language* (SPARQL).

O RDF-S é uma extensão do modelo RDF que permite a criação de hierarquias, classes, propriedade, subclasses e subpropriedades (BREITMAN *et al.*, 2003), enquanto o SPARQL é uma ferramenta que permite consultar e manipular o conteúdo publicado em RDF por meio de linguagens e protocolos especificados (SANTAREM SEGUNDO; SIMIONATO, 2016).

O uso dessas ferramentas, então, permite outros modos de representação das informações contidas na plataforma, alterando, como consequência, o relacionamento dos usuários com esse e outros ambientes digitais.

Na prática, porém, a proposta idealizada pelo consórcio W3C se mostrou mais complexa e trabalhosa do que se esperava: os esquemas atualizados da estrutura da *Web Semântica* foram sucessivamente publicados, especificamente nos anos 2000, 2002 e 2005 (RAMALHO, 2006). Nowack (2009) apresenta uma estrutura mais profunda dimensionada em três partes, na figura 1, onde se identificam as ferramentas e soluções já desenvolvidas em relação aos conceitos e abstrações da proposta.

**Figura 1** - Arquitetura de ferramentas da proposta *Web Semântica*



Fonte: Nowack (2009)

A estrutura da plataforma *Web*, os formatos de apresentação dos dados, os modelos de intercâmbio de informações e os meios de consulta e de formulação de regras são os conceitos desenvolvidos pela proposta até o momento. A figura 1 ainda destaca o segmento denominado *Linked Data*, termo apresentado por Berners-Lee em 2006 que denomina a prática de ligação (*link*) ou relacionamento entre dados e que se refere especificamente a um conjunto específico de procedimentos e ferramentas para a criação da *Web* de dados relacionados – ou ligados (BIZER; HEATH; BERNERS-LEE, 2009).

O consórcio esclarece que são esses relacionamentos que provém parte da semântica almejada pela proposta *Web Semântica* (WORLD WIDE WEB CONSORTIUM, 2015, não paginado, grifo dos autores, tradução nossa<sup>6</sup>), de forma que os dados e suas ligações possam ser interpretados pelas máquinas:

[...] a *Web Semântica* não precisa apenas de acesso aos dados, mas também **aos relacionamentos entre esses dados**, para a criação da *Web* de dados (em oposição a uma coleção pura de conjuntos de dados). Essas coleções de *datasets* interrelacionados na *Web* também podem ser denominadas *Linked Data*.

Nessa perspectiva, é importante ressaltar que *Linked Data* não é uma evolução ou superação da proposta *Web Semântica*, mas sim, um segmento interno de princípios primordiais de aplicação das ferramentas supracitadas para o alcance dos objetivos da proposta e que, no âmbito da Ciência da Informação, permite novos modos de representação de coleções de instituições *LAM* e conseqüentemente, novas formas de exploração de seus catálogos, como previamente apontado por Marcondes (2016).

Ao mesmo tempo, as discussões sobre representação de dados no ambiente digital, dentre eles a *Web*, evidenciam questões relativas a dados e metadados - objetos de estudo e de aplicação comuns aos profissionais da informação nos últimos dois séculos (GILLILAND, 2016). Esse capítulo, portanto, apresenta os conceitos e discussões sobre dados e metadados, os princípios *Linked Data*, o movimento *Linked Open Data* e as ferramentas envolvidas nos processos de descrição de coleções de patrimônio cultural no escopo desse movimento.

---

<sup>6</sup> “Furthermore, not only does the Semantic Web need access to data, but relationships among data should be made available, too, to create a Web of Data (as opposed to a sheer collection of datasets). This collection of interrelated datasets on the Web can also be referred to as *Linked Data*”

### 3.1 Dados e Metadados

Organizar, representar e recuperar informações são atividades primordiais da Ciência da Informação que não devem ser confundidas entre si. Glushko (2013) diferencia a organização de objetos da organização de informações sobre objetos, esta última denominada representação, e explica a relação inversa entre esta e a recuperação: quanto maior o esforço em representar, menos é necessário para recuperar. Embora se relacionem, Glushko (2013) explica que é fundamental que essas atividades sejam delineadas e executadas de forma isolada, uma vez que no ambiente digital a tendência é a execução das três atividades em uma única plataforma.

Em se tratando do patrimônio cultural, presente em bibliotecas, arquivos, museus e galerias de arte, Simionato, Arakaki e Santos (2017) resumem que até então os processos de representação em cada um desses ambientes buscavam atender a necessidades específicas de suas tipologias de acervos, de modo que o gerenciamento de metadados por meio das atuais tecnologias fornecem modos comuns e interoperáveis de acesso, uso e reuso de recursos. Assim, os ambientes *LAM* podem desenvolver procedimentos de representação que promovam a integração dos dados gerados sem descartar as particularidades de cada tipo de acervo.

O termo objeto informacional, ou recurso informacional, tem sido utilizado em detrimento de documento quando se fala em representação da informação - especialmente no ambiente digital, e pode ser definido como a “[...] informação objetivada no contexto de um campo do conhecimento, podendo ser apresentada em uma estrutura analógica e/ou digital. (SANTOS; SIMIONATO; ARAKAKI, 2014, p.148). Gilliland (2016), por sua vez, caracteriza o recurso informacional como qualquer objeto que possa ser localizado e manipulado como uma entidade discreta por humanos ou sistemas de informação, sendo constituído de conteúdo, contexto e estrutura.

A representação de recursos, mesmo nos meios analógicos, é conceituada como a definição de quais características, ou atributos, são necessários para a descrição desse recurso. Estes atributos recebem a denominação de metadados:

Metadados são atributos que representam uma entidade (objeto do mundo real) em um sistema de informação. Em outras palavras, são elementos descritivos ou atributos referenciais codificados que representam características próprias ou atribuídas às entidades; são ainda dados que descrevem outros dados em um sistema de informação, com o intuito de



identificar de forma única uma entidade (recurso informacional) para posterior recuperação (ALVES, 2010, p.47).

Ao afirmar que os atributos podem ser próprios ou atribuídos aos recursos, Alves (2010) evidencia as características apontadas por Gilliland (2016), que explica que o conteúdo de um recurso é intrínseco a ele, que seu contexto é extrínseco e que sua estrutura pode ter ambas as propriedades.

Glushko (2013) e Gill (2016) introduzem a ideia de estrutura, mas relacionada ao processo de representação, quando afirmam que metadados são “[...] uma definição estruturada dos atributos essenciais de um objeto informativo” (GILL, 2016, p. 03). Dessa forma, a estruturação dos metadados difere da estrutura própria do recurso, embora as duas formas se relacionem: a estrutura do recurso pode ser especificada na sua descrição (RILEY, 2017).

A segunda definição de Alves (2010) sobre metadados como dados a respeito de outros dados gera discussões conceituais entre os dois conceitos. Para a Ciência da Informação, dados são a decomposição máxima de uma informação, sendo essa decomposição constituída de uma entidade (objeto), seu atributo e o valor deste atributo (SANTOS; SANTANA, 2013). Assim, o dado se constitui do metadado acrescido do valor do recurso. Por exemplo, se o título de uma obra é “Dom Casmurro”, a obra em si é a entidade, o atributo “título” isoladamente é o metadado, e “Dom Casmurro” o valor. Nesse sentido, Pomerantz (2015) define que metadados estão além dos dados, são declarações sobre eles.

Gilliland (2016) afirma que metadados comprovam a autenticidade e o grau de completude do recurso, estabelecem o seu contexto, identificam suas relações estruturais com outros recursos, provêm diversos pontos acesso para diferentes tipos de usuários e podem fornecer informações que são geralmente obtidas por meios tradicionais.

Os metadados possuem diferentes características, categorias e funções, que são aplicáveis de acordo com o contexto em que são utilizados (GILLILAND, 2016). Riley (2017) elenca os tipos, ou categorias de metadados, apresentados no quadro 2, com exemplos de propriedades e de uso para cada um deles.

**Quadro 2 - Tipologia dos metadados**

Tipos de metadados		Exemplos de propriedades	Usos primários
Descritivos		Título Autor Assunto Gênero Data de publicação	Descoberta Apresentação Interoperabilidade
Estruturais		Sequência Posição na hierarquia	Navegação
Linguagens de marcação		Parágrafo Cabeçalho Lista Nome Data	Navegação Interoperabilidade
Administrativos	Técnicos	Tipo do arquivo Formato do arquivo Tempo/Data de criação Esquema de compressão	Interoperabilidade Manutenção de objetos digitais Preservação
	Preservação	<i>Checksum</i> <sup>7</sup> Ação de preservação	Interoperabilidade Manutenção de objetos digitais Preservação
	Direitos autorais	<i>Status</i> do copyright Termos de licença Detentores dos direitos	Interoperabilidade Manutenção de objetos digitais

Fonte: Adaptado de Riley (2017).

Para Glushko (2013) o processo de descrição contempla as ações de identificação e caracterização da essência do recurso. Assim, os metadados descritivos são os mais comuns e difundidos nas práticas de representação. O quadro 2 ainda apresenta metadados técnicos, de preservação e de direitos autorais, que podem ser agrupados como metadados administrativos, pois dizem a respeito à manutenção do recurso e/ou sua representação ao longo do tempo (RILEY, 2017). Finalmente, metadados estruturais relacionam recursos entre si e no ambiente e as linguagens de marcação permitem a navegação e a interoperabilidade entre sistemas.

Pomerantz (2015) elenca ainda uma categoria de metadados de uso, nomeados como metadados de manutenção por Glushko (2013), que registram dados relativos aos usos de recursos por parte dos usuários - por exemplo, número de *downloads* de um arquivo em um repositório. Essa categoria possui uma função importante na Ciência da Informação, pois avalia não apenas os usos dos recursos, mas também pode verificar outros hábitos dos usuários em um determinado ambiente.

<sup>7</sup> Soma de verificação de integridade de dados. O termo é formado por “*check*” – prova, verificação – e “*sum*”, de “*summation*” – soma de dígitos. Assim, é prova por soma de dígitos. (SUCESU, [19--])

Os usos primários indicados no quadro 2 introduzem e justificam as funções exercidas por metadados: a) criação (em único ou em múltiplos formatos), contextualização e reuso de recursos, b) organização e descrição, c) validação de autenticidade e veracidade, d) busca e recuperação, e) utilização e preservação e f) indicação de disponibilidade (GILLILAND, 2016). Tradicionalmente, a Ciência da Informação tem como focos a organização, descrição, busca e recuperação de recursos em bibliotecas, arquivos e museus por meio de metadados descritivos, embora outros tipos de metadados também sejam utilizados e extrapolem o âmbito dessas instituições (POMERANTZ, 2015; RILEY, 2017).

Uma vez que estão presentes no ambiente digital, metadados exercem um papel fundamental na Sociedade da Informação: fazem parte da infraestrutura moderna, assim como a energia elétrica (POMERANTZ, 2015), e possuem *status* de *commodity* (GILL, 2008). Assim, a partir dos tipos e funções generalizados de metadados, o quadro 3 apresenta suas diferentes características relacionados aos seus atributos: fonte e modo de criação, natureza, *status*, estrutura, semântica e nível do recurso.

**Quadro 3 - Atributos e características dos metadados**

<b>Atributos</b>	<b>Características</b>
Fonte de criação	Metadados intrínsecos ao recurso Metadados internos gerados no momento da criação ou digitalização de um recurso digital
	Metadados externos, ou extrínsecos, gerados em momento posterior à criação do recurso
Modo de criação	Automático, gerados por computadores
	Manual, gerados por seres humanos
Natureza	Criados por não especialistas - leigos
	Criados por especialistas - <i>expert</i>
<i>Status</i>	Estáticos – não sofrem alteração
	Dinâmicos – sofrem alterações em longo prazo
	Dinâmicos – sofrem alterações em curto prazo
Estrutura	Metadados padronizados ou com estrutura proprietária
	Não estruturados
Semântica	Metadados com valores controlados por vocabulários, listas de autoridade etc.
	Metadados com valores não controlados
Nível	Coleção de itens
	Item individual

Fonte: Baseado de Gilliland (2016)

Gilliland (2016) caracteriza os metadados como intrínsecos ou extrínsecos aos recursos, e acrescenta também como fonte os metadados que são gerados no momento da criação (analógica ou digital) de um recurso, denominados como metadados internos. Quanto ao modo de criação, podem ser gerados automaticamente por computadores ou manualmente

por seres humanos. Estes podem ser leigos ou *experts*, o que evidencia os atributos de natureza e de semântica dos metadados – seus conteúdos podem ser controlados por vocabulários ou constituírem-se de linguagem natural. Os metadados ainda podem ser estáticos ou dinâmicos, sofrendo alterações a curto e longo prazo, e descrever um único item ou uma coleção de recursos (ALVES, 2010; GILLILAND, 2016).

Finalmente, o atributo relativo à estrutura é o foco do gerenciamento de metadados, pois estes podem ser estruturados, semiestruturados ou não estruturados (ALMEIDA, 2002). Todos os atributos são combinados de acordo com a formalidade necessária no ambiente em que são aplicados (SANTOS; ALVES, 2009). Alves (2005) distribui os metadados em formatos, de acordo com a estrutura adotada:

- Formato simples: metadados não estruturados, com semântica reduzida e recuperação automática por robôs;
- Formatos estruturados: metadados estruturados de acordo com normas emergentes;
- Formatos ricos: metadados complexos, com alto grau de descrição por meio de normas, padrões e códigos específicos.

Normas, códigos e padrões são aplicados em todos os tipos de metadados e, quando relacionados ao atributo de semântica, em seus valores. O quadro 4 apresenta os padrões que podem ser utilizados no âmbito da representação do patrimônio cultural, de acordo com a tipologia elaborada por Gilliland (2016): padrões relativos à estrutura, valores, conteúdos e intercâmbio/formato dos dados.

**Quadro 4** - Padrões de metadados para patrimônio cultural

<b>Tipologia</b>	<b>Padrões de metadados</b>
Padrões de estrutura de dados	<i>Dublin Core Metadata Element Set, VRA Core Categories (VRA Core)</i>
Padrões de valores de dados	<i>LC Thesaurus for Graphic Materials (TGM), Art &amp; Architecture Thesaurus (AAT), Getty Thesaurus of Geographic Names (TGN), The Union List of Artist Names (ULAN), The cultural objects name authority (CONA), ICONCLASS</i>
Padrões de conteúdo de dados	<i>Categories for the Description of Works of Art (CDWA), Cataloging Cultural Objects (CCO), SPECTRUM</i>
Padrões de intercâmbio e formato de dados	<i>CDWA Lite XML schema, Simple Dublin Core XML schema, Qualified Dublin Core XML schema, VRA Core 4.0 XML schema, SPECTRUM XML Schema</i>

Fonte: Elaborado pela autora.

Os padrões de estrutura, como a denominação esclarece, tratam da estruturação dos atributos que descrevem o recurso, ou seja, normalizam os metadados necessários em determinadas situações. Os padrões de estrutura utilizados no domínio do patrimônio cultural são:

- *Dublin Core Metadata Elements Set*: possui estrutura simples, compreendendo um conjunto de quinze metadados básicos para descrição de recursos - contribuidor, cobertura, criador, data, descrição, formato, identificador, língua, publicador, relações, direitos, fonte, assunto, título e tipo, conhecidos pela sigla DC. Criados pela *Dublin Core Metadata Initiative* (DCMI), esses elementos são parte de um conjunto maior de metadados padronizados e especificações técnicas para representação de recursos na *Web*, denominado *DCMI Metadata Terms* (DCTERMS) (RILEY, 2017). Embora não seja um padrão específico para patrimônio cultural, é frequentemente utilizado em bibliotecas e repositórios digitais. A iniciativa DCMI é um projeto mantido pela *Association for Information Science & Technology* (ASIS&T) (DUBLIN CORE METADATA INITIATIVE, 2018).
- *VRA Core Categories (VRA Core)*: padrão para descrição de trabalhos de arte, cultura e suas imagens digitais. É desenvolvido em parceria entre a *Visual Resources Association* e a *Library of Congress* (LIBRARY OF CONGRESS, 2014).

As tipologias de valor e de conteúdo são diferentes entre si: valores se referem aos termos utilizados por controle de autoridades e vocabulários controlados, enquanto os padrões de conteúdo delimitam regras de sintaxe e de formato para preenchimento dos campos - regras de catalogação, por exemplo.

Os vocabulários e controles de autoridade AAT, TGN, ULAN e CONA, apresentados no quadro 4, são provenientes do *The Getty Research Institute* e, por isso, denominados Vocabulários *Getty*. Suas terminologias cobrem as artes em geral, a arquitetura, as artes decorativas, os arquivos, os materiais bibliográficos, suas representações visuais, e aspectos de conservação (THE GETTY RESEARCH INSTITUTE, [20--]). Estes e os demais padrões de valores utilizados pela comunidade de patrimônio cultural são:

- *Art & Architecture Thesaurus* (AAT): vocabulário genérico com termos, descrições e outros metadados para artes, arquitetura e outros tipos de patrimônio cultural;
- *Getty Thesaurus of Geographic Names* (TGN): vocabulário com termos, descrições e outros metadados sobre cidades históricas, impérios, sítios arqueológicos e características físicas importantes do ponto de vista histórico ou arquitetônico;
- *The Cultural Objects Name Authority* (CONA): compilação de títulos, atributos, assuntos e outros metadados sobre obras de arte, arquitetura e outros tipos de patrimônio cultural vinculados a coleções, instituições e outros recursos;
- *The Union List of Artist Names* (ULAN): vocabulário para controle de autoridade que inclui nome, biografia, pessoas relacionadas e outros metadados sobre artistas, arquitetos, empresas, estúdios, museus, patronos e outras pessoas ou grupos envolvidos na criação e estudo das artes e da arquitetura;
- *LC Thesaurus for Graphic Materials* (TGM): vocabulário da *Library of Congress* para indexação de materiais visuais por assunto e gênero/formato. Possui em torno de 7000 termos de assuntos e 650 termos de gênero/formato para indexação de fotografias, impressões, desenhos gráficos e outros (LIBRARY OF CONGRESS, 2018);
- ICONCLASS: criado pelo professor de História da Arte Henri van de Waal (1910-1972), da Universidade de Leiden (Holanda) na década de 50, o ICONCLASS é um sistema de classificação de assuntos para artes e iconografia amplamente utilizado por museus e instituições de arte. Foi posteriormente desenvolvido e publicado entre 1973 e 1985 pela *Royal Netherlands Academy of Arts and Sciences* (KNAW), da qual Van de Waal foi membro. Edições *online* têm sido atualizadas e publicadas desde os anos 2000 (ICONCLASS, 2012).

Os padrões de conteúdo complementam os vocabulários e controles de autoridade na descrição de recursos. Os padrões utilizados pela comunidade de patrimônio cultural são:

- *Categories for the Description of Works of Art* (CDWA): padrão voltado para a representação descritiva de obras de arte e arquitetônicas e materiais culturais individuais ou em grupos/coleções. Permite a indexação de imagens dos recursos

representados, com o objetivo de disseminar objetos culturais. Possui um total de 540 categorias e subcategorias de descrição, sendo um pequeno subconjunto delas considerados essenciais para a representação mínima do recurso: objeto/trabalho, classificação, título/nome da obra, criador, medidas, material/técnica, assunto, localização atual (THE GETTY RESEARCH INSTITUTE, 2017a). Para Gilliland (2016) o CDWA é um padrão estrutural, enquanto o *The Getty Research Institute*, responsável pelo mesmo, o classifica como padrão de conteúdo.

- *Cataloging Cultural Objects* (CCO): patrocinado pela *Visual Resources Association Foundation*, o denominado CCO é um manual intitulado “Guia para descrição de obras culturais e suas imagens” para uso da comunidade de patrimônio cultural. Não apresenta um padrão propriamente dito, mas uma série de recomendações, entre elas o uso do CDWA e dos vocabulários *Getty* para representação de recursos artísticos e arquitetônicos, incluindo pinturas, esculturas, fotografias, construções, instalações, mídias visuais, manuscritos e suas imagens digitais (VISUAL RESOURCES ASSOCIATION, 2006a).
- SPECTRUM: padrão voltado para os processos de documentação de objetos museológicos, incluindo sua descrição e manutenção, desenvolvido pela *Museum Documentation Association* (MDA) (COLLECTIONS TRUST, 2017), organização do Reino Unido que passou a ser denominada *Collections Trust* a partir de 2008. A atualização mais recente do padrão, o Spectrum 5.0, está organizada em nove processos primários, que incluem os diferentes modos de aquisição do objeto (guarda, transferência de título, permuta, etc.), catalogação, realocação e empréstimo (para outros museus e/ou exposições) (COLLECTIONS TRUST, 2017).

Finalmente, os padrões de formato e intercâmbio de dados são os responsáveis pela integração e interoperabilidade entre os sistemas e utilizam-se amplamente das linguagens de marcação, apontadas como um tipo específico de metadados por Riley (2017). Os padrões relativos ao intercâmbio de dados de patrimônio cultural refletem os padrões de estrutura apresentados anteriormente:

- *CDWA Lite XML schema*: esquema de estruturação de registros em documentos XML a partir do padrão CDWA e das recomendações CCO (THE GETTY RESEARCH INSTITUTE, 2013) para *harvesting* de acordo com o protocolo OAI-

PMH (*Open Archives Initiative Protocol for Metadata Harvesting*) (VISUAL RESOURCES ASSOCIATION, 2006b);

- *Simple Dublin Core XML schema*: esquema para estruturação de registros em documentos XML a partir da codificação do conjunto DC (DUBLIN CORE METADATA INITIATIVE, 2017);
- *Qualified Dublin Core XML schema*: esquema para estruturação de registros em documentos XML, mas este a partir da codificação do conjunto mais amplo DCTERMS (DUBLIN CORE METADATA INITIATIVE, 2017);
- *VRA Core 4.0 XML schema*: esquema para estruturação de registros em documentos XML a partir da codificação do *VRA Core Categories* (LIBRARY OF CONGRESS, 2015).
- *SPECTRUM XML Schema*: elaborado pelo *Consortium for the Computer Interchange of Museum Information* (CIMI) para a codificação de informações descritivas de objetos museológicos, incluindo informações sobre pessoas, lugares e eventos relacionados a esses recursos, assim como informações sobre sua gestão e uso dentro dos museus (GRANT, 2002). O consórcio foi desfeito em 2003 e tinha como objetivo o desenvolvimento e a disseminação de padrões de preservação digital de informações museológicas (OCLC, 2018). Este esquema baseia-se na segunda edição da norma, que atualmente encontra-se na quinta edição, evidenciando, portanto, necessidade de eventual atualização.

O primeiro dos padrões apresentados, o *CDWA Lite*, utiliza-se do OAI-PMH, um protocolo de *harvesting* de metadados que permite a interoperabilidade entre repositórios (OPEN ARCHIVES INITIATIVE, [199-]). Nota-se igualmente que a sigla XML está presente em todos os padrões de intercâmbio elencados no quadro 4. Esta representa a linguagem XML, recomendada pelo consórcio W3C para troca de dados, seja na *Web* ou em outros ambientes digitais (WORLD WIDE WEB CONSORTIUM, 2016).

No ambiente digital, em especial no escopo da proposta *Web Semântica*, todos os aspectos de metadados discutidos são os elementos “chave” para a descrição e manutenção de recursos digitais. Ao mesmo tempo, a estruturação padronizada desses metadados é determinante para as questões de interoperabilidade, como apontado pelos padrões de formato e intercâmbio de dados.



O crescimento da plataforma *Web* e a complexidade dos objetivos da proposta *Web Semântica* demandaram o desenvolvimento de outras ferramentas para gerenciamento e intercâmbio de dados e metadados. Os procedimentos de aplicação dessas ferramentas receberam a denominação de princípios *Linked Data* e geraram o movimento *Linked Open Data*, ambos discutidos a seguir.

### 3.2 Os princípios *Linked Data* e o movimento *Linked Open Data*

Ao apresentar os princípios *Linked Data*, Berners-Lee (2006) define que a proposta *Web Semântica* se trata de criar *links*, e que esses *links*, ou relacionamentos, são os responsáveis pela semântica que permite a interpretação de dados pelas máquinas. A partir dessa afirmativa, ressaltam-se dois pontos de discussão que devem ser diferenciados para o entendimento dos princípios *Linked Data* e do movimento *Linked Open Data*: a disponibilidade de dados na *Web* e as práticas de ligação desses dados.

Uma vez que esses princípios se relacionam diretamente com a plataforma *Web*, presume-se que os dados estejam disponíveis para manipulação e uso, de acordo com os princípios *Open Data*. Neste princípio, dados abertos são dados publicados sob uma licença aberta (ARAKAKI, 2016). Tal consideração justifica a afirmativa de Heath e Bizer (2011, não paginado, tradução nossa<sup>8</sup>) de que “[...] o termo *Linked Data* refere-se a um conjunto de melhores práticas para a publicação e interligação de dados estruturados na *Web*”.

Berners-Lee (2006), porém, retifica seu texto<sup>9</sup>, explicando que isso nem sempre acontece: dados podem estar ligados, mas não necessariamente abertos. Nesse sentido, os princípios *Linked Data* referem-se especificamente aos procedimentos de ligação de dados, abertos ou não, por meio do uso sistemático de ferramentas desenvolvidas pela proposta *Web Semântica*. Estes princípios compreendem os seguintes direcionamentos:

- Utilizar URIs para nomear recursos;
- Utilizar HTTP URIs, para que pessoas possam procurar por esses nomes;
- Quando alguém procurar por um URI, fornecer informações úteis por meio do uso de padrões (RDF, SPARQL);

---

<sup>8</sup> “The term *Linked Data* refers to a set of best practices for publishing and interlinking structured data on the *Web*”

<sup>9</sup> Atualização do texto de 2006, publicado em 2010.

- Incluir *links* para outros URIs, para a descoberta de outros recursos. (BERNERS-LEE, 2006, *online*, tradução nossa<sup>10</sup>; HEATH; BIZER, 2011)

A aplicação desses princípios permite a recuperação de todos os *links* relacionados a determinado URI (BERNERS-LEE, 2006). Em outras palavras, permite a descoberta de todas as informações e relacionamentos de um recurso que esteja assim identificado.

A utilização de URI (primeiro princípio) fornece uma identificação única de um recurso, e sua associação com o protocolo *Hypertext Transfer Protocol* (HTTP), de acordo com o segundo princípio, permite que ele seja “encontrável” no ambiente *Web*. Isso em razão do HTTP ser um protocolo padrão para comunicação e localização de documentos, ou recursos, na *Web* (RAMALHO, 2006). A descrição e o relacionamento (terceiro princípio) com outros recursos são efetivados pela aplicação do modelo RDF, enquanto que sua recuperação e manipulação, enquanto dados brutos, ocorre por meio do uso do protocolo SPARQL - no desenvolvimento da *Web* de dados, estes são separados da formalidade e da camada de apresentação (BIZER; HEATH; BERNERS-LEE, 2009).

O quarto princípio, em concordância com o terceiro, é o responsável pela efetiva ligação de dados provenientes de diferentes fontes e retomam a ideia de que os benefícios da ligação de dados podem ser aproveitados de outros modos se esses dados forem abertos.

Nessa perspectiva, entende-se que princípios, de forma geral, são requisitos fundamentais que quando não observados podem comprometer os objetivos almejados. Os movimentos surgem, então, para promover e incentivar ações que reflitam os princípios. Neste caso, os princípios *Linked Data*, em consonância com os objetivos maiores da proposta *Web Semântica* e com os princípios *Open Data*, culminaram no movimento *Linked Open Data*, como ilustrado na figura 2, abaixo.

---

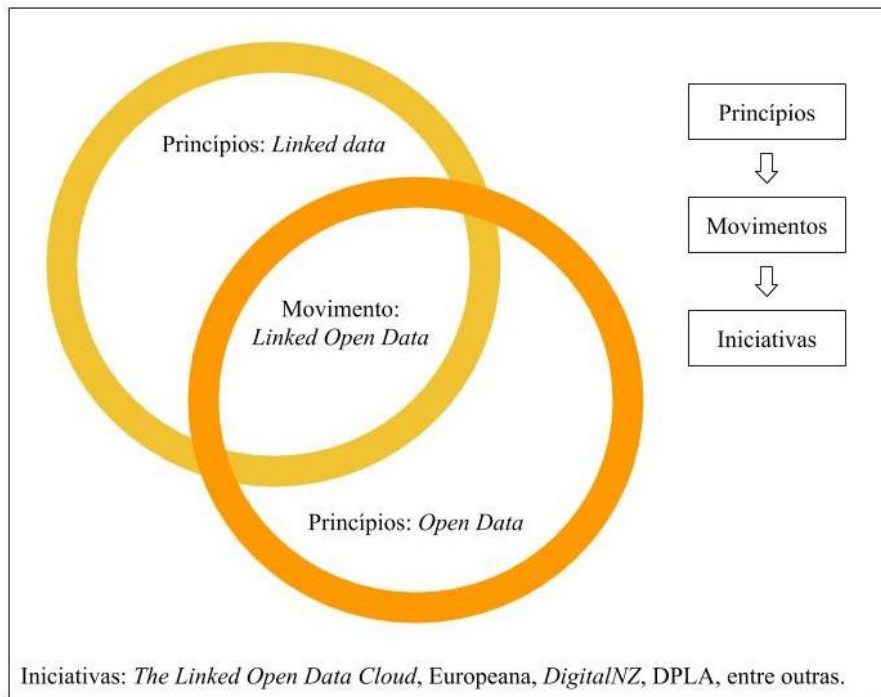
<sup>10</sup> “Use URIs as names for things

Use HTTP URIs so that people can look up those names.

When someone looks up a URI, provide useful information, using the standards (RDF\*, SPARQL)

Include links to other URIs. so that they can discover more things”

**Figura 2** - Diagrama de princípios, movimentos e iniciativas relativas ao *Linked Data*



Fonte: Elaborado pela autora.

O movimento *Linked Open Data*, então, trata da aplicação dos princípios *Linked Data* em dados abertos (BERNERS-LEE, 2006) e tem como objetivo a publicação de dados estruturados na *Web*. Desse modo, os dados publicados não estão apenas disponíveis para visualização e acesso, mas estão estruturados de modo a permitir sua efetiva manipulação, uso e reuso.

Ainda de acordo com a figura 2, percebe-se que os movimentos se traduzem e se concretizam em iniciativas, que por sua vez englobam projetos, instituições, grupos de pesquisadores e de profissionais. Assim, o movimento *Linked Open Data* tem gerado iniciativas concretas de elaboração e execução de projetos de abertura e ligação de dados de patrimônio cultural. A iniciativa *The Linked Open Data Cloud (LOD-Cloud)* será abordada abaixo, enquanto as demais serão apresentadas no próximo capítulo.

Como forma de incentivo e avaliação da qualidade dos dados publicados de acordo com o movimento *Linked Open Data*, Berners-Lee (2006, *online*, tradução nossa<sup>11</sup>) também

<sup>11</sup> “

★ Available on the web (whatever format) but with an open licence, to be Open Data

★★ Available as machine-readable structured data (e.g. excel instead of image scan of a table)

★★★ as (2) plus non-proprietary format (e.g. CSV instead of excel)

★★★★ All the above plus, Use open standards from W3C (RDF and SPARQL) to identify things, so that people can point at your stuff

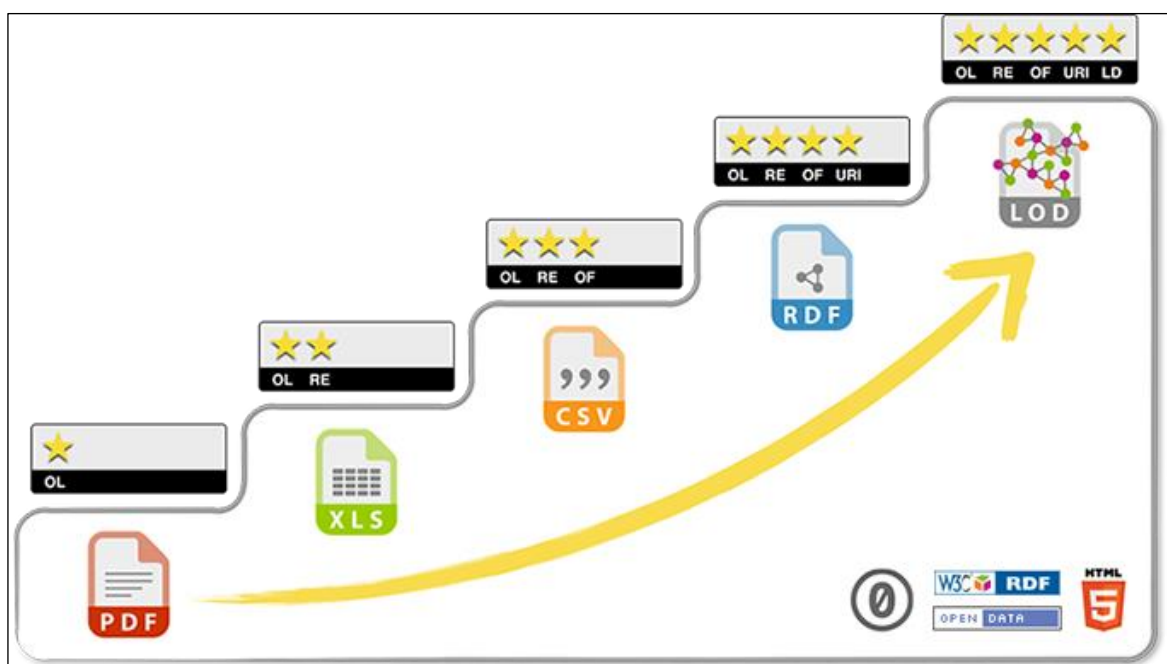
★★★★★ All the above, plus: Link your data to other people's data to provide context”

elaborou um *ranking* cinco estrelas para publicação e ligação de dados, o *5-Star Linked Open Data*:

- Uma estrela (★): dados disponíveis na *Web* (em qualquer formato) com licença aberta (*Open data*);
- Duas estrelas (★★): dados disponíveis em formatos estruturados legíveis por máquina (Formato *Excel* ao invés de uma imagem digitalizada de uma tabela, por exemplo);
- Três estrelas (★★★): publicação de dados em formatos não proprietários (Formato *CSV* ao invés de *Excel*, por exemplo);
- Quatro estrelas (★★★★): uso dos padrões abertos do W3C (URI, RDF e SPARQL) para identificação e descrição de recursos, para que pessoas possam referenciá-los;
- Cinco estrelas (★★★★★): relacionar seus dados com dados de outras pessoas para contextualização.

De acordo com esse *ranking*, as primeiras três pontuações dizem respeito aos princípios *Open Data*, uma vez que os dados são publicados com licenças abertas e em formatos estruturados que permitem sua manipulação e uso. As pontuações ‘quatro’ e ‘cinco estrelas’, porém, tratam dos princípios *Linked Data*, pois enriquecem a publicação por meio da descrição e ligação desses dados, permitindo seu relacionamento com outros dados, ou conjuntos de dados – denominados *datasets*. A figura 3 apresenta a visualização do *ranking* de acordo com as exigências de cada pontuação.

**Figura 3 - Ranking “5 estrelas” para publicação e ligação de dados**



Fonte: Kim e Hausenblas (2015).

Por meio da observação de cada uma das pontuações, percebe-se que essa classificação é cumulativa, ou seja, para que um determinado conjunto de dados seja considerado “cinco estrelas” ele deve estar de acordo com todos os requisitos anteriores.

Bizer, Heath e Berners-Lee (2009) lembram que na *Web* de Dados qualquer pessoa pode publicá-los, que dados podem ser de qualquer tipo e que, por isso, a *Web* é genérica. Assim, diversos tipos de *datasets* podem ser encontrados atualmente, publicados por indivíduos ou por instituições, com diversos objetivos e domínios de especialidade. Destacam-se as publicações de dados governamentais e científicos, impulsionados pelos princípios *Open Data* e por questões de transparência pública.

O *LOD-Cloud* é um exemplo desse tipo de iniciativa, alavancado por um grupo de trabalho do consócio W3C denominado *Semantic Web Education and Outreach Interest Group (SweoIG)*. Embora encerrado em 2008, o grupo tinha como objetivo conscientizar a comunidade acerca dos benefícios de uma *Web* com estrutura semântica e promover a educação e o uso das soluções e tecnologias desenvolvidas para a proposta *Web Semântica* (WORLD WIDE WEB CONSORTIUM, 2008). O consócio esclarece que embora o grupo não esteja mais atuando, algumas de suas atividades mantêm-se operantes pela comunidade, dentre elas, o *LOD-Cloud*.

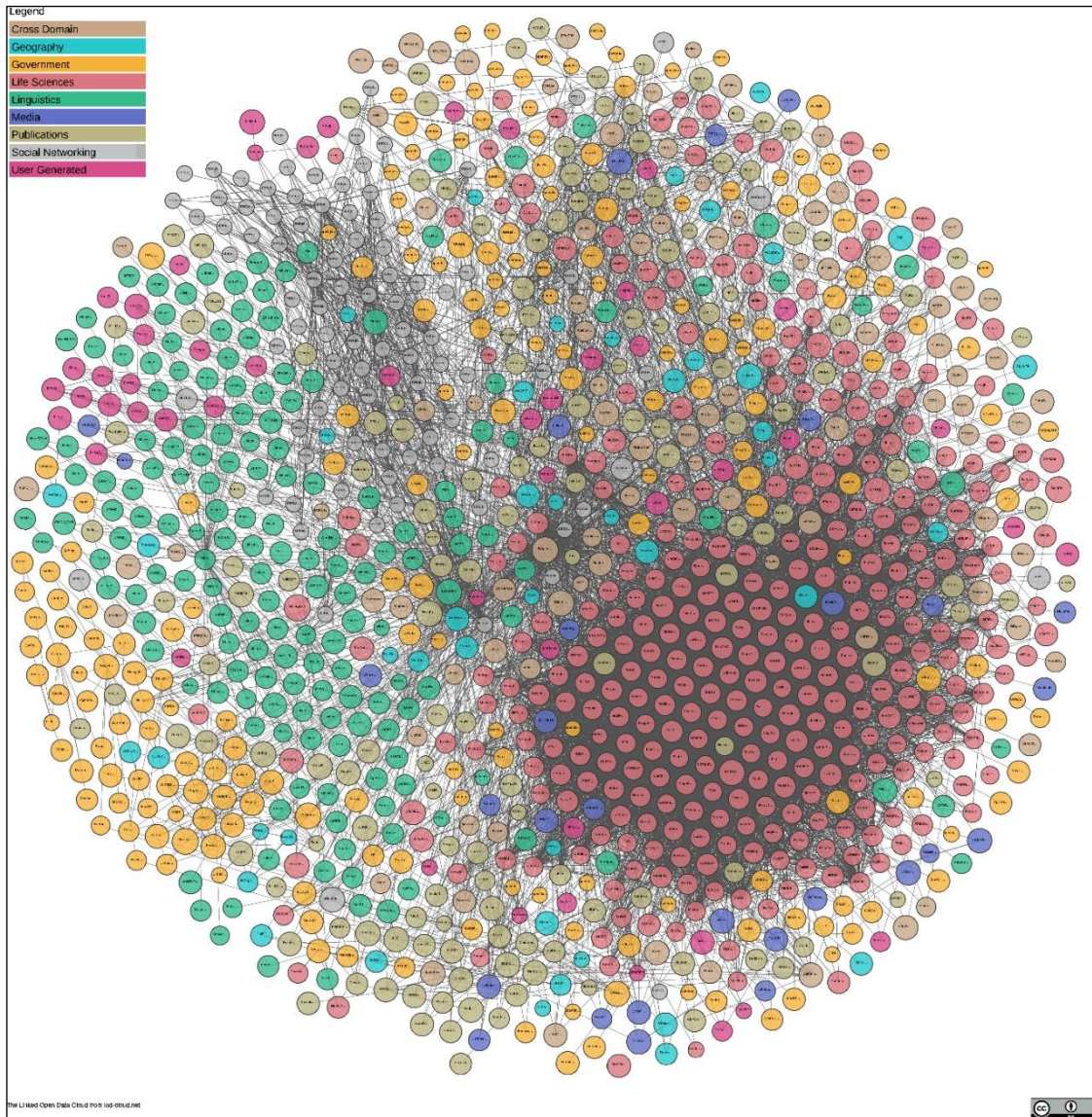
Fundado em janeiro de 2007, o projeto é uma iniciativa concreta de conversão, quando necessário, e publicação de dados (BIZER; HEATH; BERNERS-LEE, 2009), bem como de criação de uma rede de *datasets* conectados em concordância aos parâmetros do movimento *Linked Open Data*. A figura 4<sup>12</sup>, a seguir, apresenta a visualização de todos os *datasets* atuais do projeto, seus relacionamentos e domínios.

---

<sup>12</sup> A figura demonstra o *LOD-Cloud* em sua última atualização, no dia 30 de maio de 2018.



**Figura 4** - Visualização de *datasets* do projeto *LOD-Cloud*



Fonte: McCrae *et al.* (2018).

O número de *datasets* contabilizados a cada atualização da rede demonstra o rápido crescimento do projeto nos seus onze anos de existência, uma vez que inicialmente contava-se com doze conjuntos de dados publicados, passando para 570 em 2014, e alcançando a contagem de 1186 conjuntos de dados publicados até maio de 2018 (MCCRAE *et al.*, 2018). Estes *datasets* estão distribuídos pelos domínios: publicações em geral, ciências da vida, domínio geral (*cross-domain*), dados geográficos, dados governamentais, mídia, dados de uso geral, dados de redes sociais e linguísticos.

Dentre os conjuntos de dados publicados destaca-se, por exemplo, o *dataset DBpedia* – conjunto de dados da *Wikipedia* – uma vez que se constitui de dados de diferentes domínios e, por isso, possui alto número de ligações com outros *datasets*. No escopo do projeto encontram-se ainda *datasets* relativos a coleções de patrimônio cultural publicados por

instituições e iniciativas que identificaram oportunidades e benefícios no uso das tecnologias para a representação de recursos, bem como na disponibilização de dados para o acesso às suas coleções. Tais iniciativas serão abordadas no capítulo 4.

A partir, então, dos princípios *Linked Data*, das diretrizes do movimento *Linked Open Data* e dos padrões de metadados desenvolvidos e adotados por comunidades que trabalham com patrimônio cultural, apresentam-se as ferramentas desenvolvidas para a descrição de recursos no ambiente *Web* que são utilizadas para a representação de coleções culturais.

### 3.3 *Linked Open Data* e a descrição do patrimônio cultural

Considerando-se o extenso uso da linguagem XML pelas comunidades que desenvolveram os padrões de intercâmbio de dados de patrimônio cultural, bem como as ferramentas do movimento *Linked Open Data* que se relacionam com a descrição de recursos nos ambientes digitais, esta subseção discute brevemente as ferramentas URI, XML e RDF, previamente apresentadas no início do capítulo.

Recursos possuem identidade e, por isso, necessitam de uma identificação que reflita e assegure a sua persistência, efetividade, autenticidade e proveniência ao longo do tempo (GLUSHKO, 2013). No contexto da *Web*, a identificação de recursos é realizada pelo URI, que consiste em uma atribuição significativa aos recursos, pois oferece persistência do referido nome ao seu endereço na *Web*.

Isotani e Bittencourt (2015) explicam que URIs podem ser divididos entre *Unified Resource Name* (URN), que especificam apenas o nome do recurso, e *Unified Resource Locator* (URL) que fornecem seu endereço de localização por meio de um protocolo. Estes autores ainda tratam brevemente do *International Resource Identifier* (IRI), uma forma de identificação mais ampla, pois inclui os caracteres chineses, japoneses, cirílicos e coreanos.

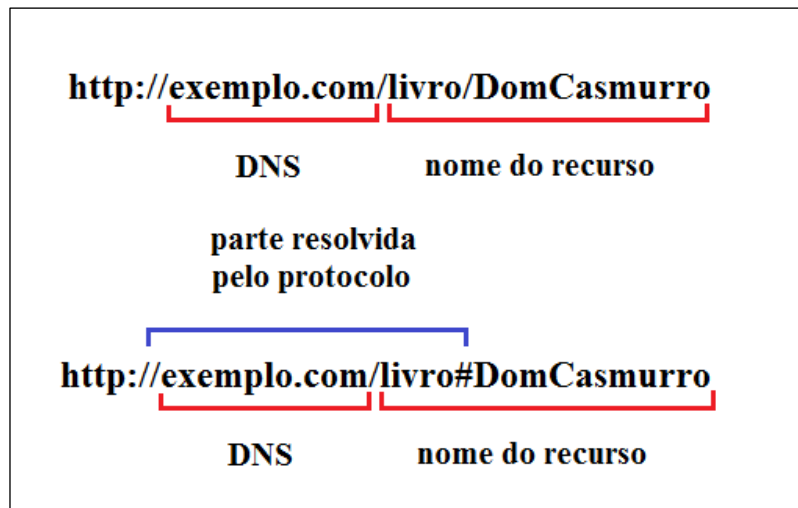
Os identificadores de um objeto real e de sua representação digital devem ser diferenciados, razão pela qual existem dois modos de criação de URIs:

- 303 URIs: este tipo redireciona o primeiro identificador, relativo ao objeto real, para o URI de sua representação digital;
- *Hash* URIs: possuem uma parte especial da identificação separada por um *hash* (#), que especifica o objeto. Dessa forma, podem-se diferenciar objetos e/ou suas

representações a partir de um mesmo identificador inicial (HEATH; BIZER, 2011).

A figura 5 apresenta exemplos dos dois tipos de URI.

**Figura 5** - Exemplos de 303 URI e *hash* URI



Fonte: Elaborado pela autora.

O protocolo padrão para localização de documentos ou recursos na *Web* é o protocolo HTTP, conceituado como “[...] um conjunto de regras definidas para a comunicação entre os componentes da *Web*” (RAMALHO, 2006, p.32). Conforme mostra a figura 5, é comum o uso de HTTP URIs, que identificam e recuperam objeto de forma unificada (HEATH; BIZER, 2011) para usuários humanos (via páginas *Web*) ou máquinas - aplicações de *software* (via dados brutos) (RILEY, 2017).

No primeiro exemplo, um URI do tipo 303 identifica o livro “Dom Casmurro”. Se este URI não se tratar de um objeto digital, o protocolo será redirecionado a um URI semelhante (mas não exatamente igual) que contém a representação desse livro no ambiente digital. Igualmente, são dereferenciados por dois protocolos HTTP (HEATH; BIZER, 2011).

O segundo exemplo também identifica o livro “Dom Casmurro”, mas utilizando um *hash* URI. Desse modo, se a obra possui uma representação digital, ou se é parte de uma coleção de livros com autoria de “Machado de Assis”, os demais itens podem receber o mesmo identificador, trocando-se apenas o título da obra após o *hash* ou a denominação do objeto para especificar que se trata de uma representação digital do mesmo.

Os dois exemplos mostram que identificadores são elaborados em linguagem alfabética: caracteres numéricos não são recomendados (WOOD *et al.*, 2014). Também não se



recomenda o uso de *namespaces* (DNS) que não sejam de controle (manipulação) do responsável pelo identificador, e que a identificação deixe evidentes características de implementação do ambiente onde os URIs estão sendo alocados (HEATH; BIZER, 2011). Tal exposição pode comprometer tanto o direcionamento do URI, quanto a integridade do ambiente.

A escolha de uso de cada um dos tipos naturalmente está relacionada às necessidades da situação, embora Wood (*et al.*, 2014) não recomende o uso de *hash* URIs em ambientes *Linked Open Data*, pois nesse tipo de identificação apenas parte do endereço (o que estiver antes do *hash*) é resolvido pelo *browser*, enquanto Heath e Bizer (2011) o recomendam para a construção de vocabulários, pois são identificadores que possuem características hierárquicas.

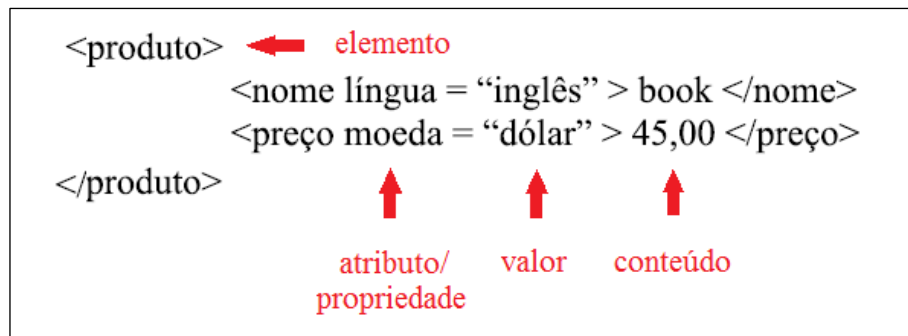
A partir de uma identificação que atenda aos requisitos elencados por Glushko (2013) pode-se trabalhar com a descrição dos recursos a partir da linguagem XML e do modelo RDF.

A linguagem XML foi originalmente desenvolvida pelo consórcio W3C a partir do *Standard Generalized Markup Language* (SGML) (ISO 8879) para otimizar a publicação de documentos em larga escala na *Web* (WORLD WIDE WEB CONSORTIUM, 2016). Isso porque a linguagem padrão de desenvolvimento de páginas Web é a *HyperText Markup Language* (HTML), uma linguagem de marcação de documentos (também proveniente do SGML) que prioriza a apresentação do conteúdo e pouco infere sobre sua representação, ou descrição semântica (ALMEIDA, 2002).

Para Ferreira (2014) o XML não é uma linguagem de marcação, mas uma ferramenta que apenas define regras e restrições sintáticas para a marcação de documentos, enquanto que Riley (2017) a considera uma multilinguagem. Apesar das divergências conceituais, XML atualmente é um importante instrumento de intercâmbio de dados da *Web* e em outros ambientes digitais (WORLD WIDE WEB CONSORTIUM, 2016).

Sua sintaxe é adequada para uso em dados semiestruturados, que se utilizam da teoria dos grafos como modelo de descrição (ALMEIDA, 2002). Um grafo (G) é composto de vértices/nós (V) e arestas (A) que relacionam dois ou mais nós (FEOFILOFF; KOHAYAKAWA; WAKABAYASHI, 2017). Dentre seus componentes básicos, portanto, estão os elementos e os atributos que os descrevem, conforme a figura 6, abaixo.

**Figura 6** - Fragmento de um documento XML



Fonte: Adaptado de Almeida (2002).

Os elementos são os blocos do documento XML (RAY, 2001) intercalados pelos sinais “<” e “>” (ALMEIDA, 2002). O exemplo da figura 6 apresenta um elemento produto e os elementos nome e preço desse produto. Os atributos, ou propriedades, de um elemento são opcionais e compostos de pares ‘nome-valor’, sendo que o valor deve estar entre aspas (ALMEIDA, 2002). No exemplo acima, o elemento ‘nome’ possui o atributo ‘língua’ com valor “inglês”, e o elemento ‘preço’ possui o atributo ‘moeda’ com o valor “dólar”.

Um documento XML apresenta ainda outros componentes. Ray (2001), Assumpção (2013) e Arakaki (2016) assemelham os seguintes apontamentos:

- Textos: são os caracteres que compõem o conteúdo de um elemento. No exemplo da figura 6, são as expressões ‘book’ e ‘45,00’;
- Comentários: anotações no documento que não são processadas por máquina;
- Instruções de processamento: orientações para processamento do documento em determinada aplicação;
- *Namespaces*: representam grupos de elementos e atributos que fazem parte de determinado vocabulário. Os prefixos de cada *namespace* são adicionados ao documento para indicar os vocabulários utilizados;
- Raiz: ponto abstrato acima do elemento que abrange todo o documento.

Desse modo, um documento XML possui estrutura hierárquica em forma de árvore (ASSUMPCÃO, 2013) e também é estruturado em grafos, onde os ‘nós’ são os elementos e as arestas são os valores que os relacionam (ALMEIDA, 2002). Os elementos e a raiz de um documento XML são os únicos tipos de nós que podem conter outros nós (RAY, 2001).

Outro recurso primordial para XML são os *XML Schema*, que atualmente substituem o *Data Type Definition* (DTD). Este é um arquivo que confere a estrutura sintática de um

documento XML e ainda permite referenciar dados externos ao documento por meio de uma URL, fato que possibilita o intercâmbio de dados (ALMEIDA, 2002). *XML Schema* mantém essa funcionalidade, mas é uma tecnologia avançada em relação ao DTD, uma vez que “define os elementos que formam um documento válido naquele formato, junto dos atributos que podem ser, em que ordem pode aparecer e quantas vezes podem ocorrer” (RILEY, 2017, p.16, tradução nossa<sup>13</sup>).

Os padrões de intercâmbio de dados de patrimônio cultural apresentados anteriormente são exemplos de *XML Schema* para conversão de registros de cada padrão de estrutura para documentos XML. Isso porque, em conjunto com o modelo RDF, o uso da linguagem XML é fundamental para a descrição de recursos no ambiente *Web*.

O RDF é definido pelo consórcio W3C como um modelo aberto de descrição de recursos na *Web*, sendo que tais recursos podem estar inseridos na plataforma ou não (FERREIRA; SANTOS, 2013). Esta diferenciação entre recursos que estão ‘dentro’ ou ‘fora’ da *Web* retoma as considerações supracitadas de URIs, que permitem a identificação de um recurso, e também de sua representação digital, evidenciando a coesão entre a identificação e a descrição de um recurso digital ou não digital.

As descrições dos recursos ocorrem a partir de triplas ‘sujeito, predicado e objeto’ (WORLD WIDE WEB CONSORTIUM, 2014), também denominadas como ‘recurso, propriedade e valor’ ou ainda, ‘entidade, atributo, valor’, que provém arquitetura baseada em grafos (BIZER; HEATH; BERNERS-LEE, 2009). Em consonância com o uso do termo recurso informacional pela Ciência da Informação, este trabalho adota a denominação “recurso, propriedade e valor” para os componentes das triplas RDF.

Tais triplas podem ser literais ou *links* RDF. Como explicam Heath e Bizer (2011), triplas literais descrevem propriedades de um recurso e seu valor é um literal (não se constitui de um URI), enquanto que *links* RDF relacionam dois recursos e, portanto, constitui-se de três URIs, o terceiro referindo-se à propriedade do recurso.

Os recursos relacionados por *links* RDF podem pertencer a uma mesma fonte ou não. Assim, Heath e Bizer (2011) subdividem os *links* RDF entre internos, que relacionam recursos de uma mesma fonte, e externos, que ligam recursos de fontes diferentes e podem ser de diferentes tipos, destacando-se os *links* RDF do tipo identidade, relacionamento e vocabulários:

---

<sup>13</sup> “An XML Schema defines the elements that make up a valid document in that format, along with the attributes each element can take, in what order they can appear, and how many times they can appear”.

- Identidade: em ambientes colaborativos e distribuídos como a *Web*, diferentes URIs podem ser atribuídos ao mesmo recurso. Esse tipo de *link*, então, relaciona esses URIs para identificá-los como referentes ao mesmo recurso e para agrupar as diferentes informações que cada URI contém sobre ele;
- Relacionamento: esses *links* relacionam recursos de diferentes fontes, sendo o primeiro o recurso e o segundo o valor da tripla;
- Vocabulários: *links* que relacionam recursos com os termos de vocabulários externos que os descrevem (HEATH; BIZER, 2011).

Gill (2016) explica que os vocabulários fornecem a semântica da descrição, uma vez que o modelo RDF provê apenas sua sintaxe, ou sua estrutura. O uso de vocabulários é fundamental para a descrição de recursos, seja na *Web* ou em outros ambientes. Como supracitado, a comunidade de patrimônio cultural desenvolveu padrões de valores de metadados, relativos aos vocabulários que descrevem esse tipo de recurso. Esses padrões também se inserem no escopo do movimento *Linked Open Data* e serão discutidos do capítulo 5.

Catarino e Souza (2012) lembram que o RDF é um modelo abstrato; assim, sua aplicação (serialização) é variada. O consórcio W3C desenvolveu duas formas de serialização padrão: o RDF/XML e o RDFa (HEATH; BIZER, 2011). O quadro 5, abaixo, apresenta estas e outras formas de serialização utilizadas na *Web*, bem como suas funções.

**Quadro 5** - Serializações RDF e suas funções

<b>Serialização RDF</b>	<b>Tipo de código</b>	<b>Quando usar</b>
RDFa	Código RDF embutido em HTML	SEO <sup>14</sup>
RDF/XML	Código RDF com estrutura em XML	Aplicações que usam estrutura XML
JSON-LD	Código RDF com estrutura JSON	Aplicações que usam JSON
<i>N-Triples</i>	Código RDF com estrutura de Triplas	Processamento e intercâmbio de <i>Big Data</i> em RDF.
<i>Turtle</i>	Código RDF para facilitar a leitura humana	Processamento e intercâmbio de <i>Big Data</i> em RDF.
<i>TriG</i>	Código com estrutura <i>Turtle</i>	Representação de múltiplos grafos
<i>N-Quads</i>	Código RDF com estrutura de Triplas	Processamento e Intercâmbio de grandes catálogos de dados.

Fonte: Isotani e Bittencourt (2015, *online*)

Wood (*et al.*, 2014) explica as principais serializações apresentadas no quadro 5, acima: RDFa tem como objetivo melhorar a busca e recuperação de dados embutidos em documentos HTML – Hypertext markup Language, enquanto que o RDF/XML é uma forma de estruturação de dados utilizando a linguagem XML; o formato *N-Triple* apresenta a estrutura RDF em forma de triplas, e o formato *Turtle* é voltado para a leitura humana, pois tem uma visualização mais simples.

O formato JSON-LD, por sua vez, é voltado para desenvolvedores que se utilizam do formato *JavaScript Object Notation* (JSON) e destaca-se por se tratar de um formato que trabalha diretamente com os princípios *Linked Data*. Mitchell (2013, p.36, tradução nossa<sup>15</sup>) explica que o formato JSON-LD

[...] é uma extensão do JSON que inclui um método para identificar dados através de IRIs, suporta a desambiguação de objetos JSON ao combinar conjuntos de dados de documentos diferentes, fornece um método para identificar linguagens e tipos de dados literais e para expressar relacionamentos gráficos usando estruturas JSON.

Isotani e Bittencourt (2015) indicam ainda, no quadro 5, as situações em que o JSON-LD, e as demais serializações são adequada para uso. Destacam-se os formatos *N-Triples*,

<sup>14</sup> *Search Engine Optimization*

<sup>15</sup> *JSON-LD is an extension of JSON that includes a method for identifying data through IRIs, supports disambiguation of JSON objects when combining datasets from different documents, provides a method for identifying language and data types of literals, and provides a method for expressing graph relationships using JSON structures.*

*Turtle* e *N-Quads*, que lidam com grandes volumes de dados, bem como o formato *TriG*, que trata de múltiplos grafos. A variedade de serializações desenvolvidas para uso do RDF demonstra, então, a importância e a versatilidade desse modelo. Dentre todas as possibilidades apresentadas, ressalta-se o RDF/XML, uma vez que os padrões de intercâmbio de dados de patrimônio cultural apresentados anteriormente baseiam-se em *XML Schema*.

Desse modo, percebe-se que a comunidade de patrimônio cultural possui direcionamentos e iniciativas para descrição de recursos com os objetivos de promoção da interoperabilidade entre sistemas e de inserção no movimento *Linked Open Data*. Os próximos capítulos, então, apresentam e discutem iniciativas institucionais de descrição e disponibilização de dados de patrimônio cultural, os padrões de metadados e os modelos de descrição que foram desenvolvidos pela comunidade de patrimônio cultural no escopo do movimento *Linked Open Data*.

#### **4 SUSTENTABILIDADE DE DADOS E INICIATIVAS *LINKED OPEN DATA* EM INSTITUIÇÕES DE PATRIMÔNIO CULTURAL**

As teorias e práticas que englobam o campo das Humanidades Digitais têm como uma de suas finalidades o uso das tecnologias para a produção de conhecimentos em Ciências Humanas e Sociais, bem como a apropriação desses conhecimentos pelos diferentes atores da sociedade. Tal apropriação é possível por meio da disponibilização de informações, inclusive culturais, de modo que se permita seu uso e reuso em diferentes contextos e ambientes.

Assim, centros de documentação e memória – que incluem instituições *LAM* - encontram nas tecnologias novos modos de promoção do acesso às suas coleções, reconhecidas pela literatura como patrimônio cultural. Ao mesmo tempo, tais instituições percebem novas possibilidades de preservação desse patrimônio: o supracitado ‘Manifesto para a Preservação Digital’, da UNESCO, considera o papel essencial de instituições *LAM* na salvaguarda e acesso ao patrimônio cultural, seja ele nativo digital ou digitalizado. Este manifesto também expressa o protagonismo que as tecnologias assumem no gerenciamento de recursos e dos dados desses recursos nos ambientes digitais.

Para cumprir, então, os objetivos propostos pela UNESCO e, de modo geral, pelo campo das Humanidades Digitais, instituições *LAM* têm buscado formas de gerenciar suas coleções no ambiente digital, com especial atenção às questões de preservação. Dentre elas, destacam-se primeiramente os processos de digitalização de acervos.

Dahlström, Hansson e Kjellman (2012) explicam que, nesses processos, as decisões tomadas influenciam o modo como o recurso é exposto para o público e, conseqüentemente, o uso deste pela comunidade: a produção de imagens relativas a determinado item, além de ser uma representação em si, atua como mediadora entre esse item e o usuário. Nesse sentido, Maron, Yun e Pickle (2013) esclarecem que a preservação não envolve apenas os conteúdos das coleções, mas também tratam do acesso e dos usos possíveis pelos usuários, bem como a geração de impacto na comunidade. Assim, pondera-se que os recursos que são nativos digitais necessitam de ações específicas para a promoção do seu acesso e uso, de modo que instituições *LAM* devem conceber processos de preservação que incluam também os procedimentos de representação de suas coleções.

De modo amplo, a representação envolve escolhas tomadas pela instituição no que diz respeito aos procedimentos de descrição e disponibilização de recursos - que incluem os produtos da sua digitalização - em que se destacam o uso de metadados e de padrões de metadados. As questões de preservação, então, envolvem todos os tipos, características e

funções de metadados, abordados por Gilliland (2016), Pomerantz (2015) e Riley (2017). Nesse contexto, Lee (*et al.*, 2017) determina que metadados são essenciais para os processos de curadoria digital, especialmente em ambientes *Open Data*. Ao mesmo tempo, a evolução dos registros informacionais em bibliotecas, arquivos e museus trata da desvinculação de seus dados dos catálogos para redes abertas e amplas, pois a publicação de dados torna-os independentes dos sistemas de gerenciamento de catálogos e ao mesmo tempo permite um novo paradigma de interligação de coleções no ambiente digital (MARCONDES, 2016).

Concomitantemente, o *Council on Library and Information Resources* (2001) considera que os programas digitais que adotam padrões e melhores práticas de trabalho, arquiteturas digitais comuns e coerentes e meios práticos de criação e compartilhamento de conhecimentos são programas voltados à sustentabilidade.

Maron, Yun e Pickle (2013, p.12, tradução nossa<sup>16</sup>) definem a sustentabilidade no âmbito da preservação digital como “[...] a habilidade de identificar os recursos necessários para manter o valor de um objeto digital ou de um serviço para os usuários aos quais se destinam”. Em outras palavras, relaciona-se a ações que mantêm o valor e a relevância de um recurso para o público. Nesse sentido, Ninin e Simionato (2017) afirmam que as decisões tomadas pelas instituições *LAM* influenciam o acesso, a disseminação e o uso dos conteúdos pelos usuários e exigem medidas da instituição para a garantia de sustentabilidade de seus produtos e serviços.

Percebe-se, então, que as ações de preservação tratam dos suportes, dos conteúdos e dos processos de representação de recursos, que geram produtos e serviços diretamente relacionados aos ambientes em que estão inseridos. As instituições *LAM*, portanto, têm buscado alcançar processos sustentáveis em todas as etapas do ciclo de gerenciamento informacional de suas coleções (NININ; SIMIONATO, 2017).

Uma vez que os catálogos são os responsáveis por expressarem os resultados desse gerenciamento, especialmente no que diz respeito à representação, recuperação e acesso aos recursos - sejam eles nativos digitais ou digitalizados - retomam-se as considerações de Gilliland (2016) sobre o conteúdo, contexto e estrutura de recursos e de Marcondes (2016) a respeito da busca de interoperabilidade entre sistemas de instituições *LAM*. Isso porque, no ambiente aberto e distribuído da *Web*, as questões de interoperabilidade têm ganhado destaque, por permitirem a ampliação das possibilidades de acesso e uso de recursos.

---

<sup>16</sup> “*The ability to identify the resources needed to keep a digital resource or service delivering value to the users it is intended to serve*”



No mesmo sentido, Dodebei (2006) discorre sobre a necessidade de se fazer os patrimônios digitais conversarem entre si. Em suas afirmativas, Dodebei (2006), Gilliland (2016) e Marcondes (2016) tratam de questões de interoperabilidade, embora em diferentes níveis: Dodebei (2006) e Gilliland (2016) focam a contextualização de recursos, o que implica questões de interoperabilidade semântica, enquanto Marcondes (2016) discorre sobre interoperabilidade técnica entre sistemas.

Nesse sentido, destacam-se as ferramentas de descrição utilizadas pelo movimento *Linked Open Data*, pois têm como objetivo a criação de relacionamentos entre recursos sob as duas perspectivas: permitem a identificação e o acesso persistente aos recursos, bem como sua contextualização em relação a outros. Marcondes (2016), então, propõe o estabelecimento de relações culturalmente significativas entre coleções culturais, utilizando-se dessas ferramentas nos processos de representação.

No escopo maior dos princípios *Open Data* e do movimento geral de publicação de dados, o consórcio W3C, por meio do grupo de trabalho *Data on the Web Best Practices Working Group*, publicou o documento ‘*Data on the Web best practices*’, em que são elencados um conjunto de melhores práticas para a publicação de dados, ligados ou não, que tem como objetivo garantir o reuso, a compreensão, a possibilidade de ligação, a descoberta, a confiança, o acesso, a interoperabilidade e o processamento dos dados publicados por consumidores e publicadores, em diferentes situações e com visões de longo prazo (WORLD WIDE WEB CONSORTIUM, 2017). Visões de longo prazo estão diretamente relacionadas às questões de sustentabilidade, ao mesmo tempo em que as questões de interoperabilidade são essenciais para a *Web* e também para ambientes *LAM*.

A aplicação de ferramentas do movimento *Linked Open Data* para a representação e disponibilização de dados de patrimônio cultural torna-se uma ação promissora de instituições *LAM* para permitir o acesso e a preservação sustentáveis de suas coleções. Nesse sentido, a literatura científica apresenta contribuições à temática de utilização dessas ferramentas para ligação e publicação de dados de patrimônio cultural.

Hyvönen (2012) contextualiza o movimento de publicação de dados em relação aos benefícios que esse movimento pode proporcionar às instituições *LAM* e às Humanidades Digitais de modo geral. Sua abordagem ultrapassa o movimento *Linked Open Data* e trata da proposta *Web Semântica* com um todo, discutindo os sistemas semânticos e seus componentes, bem como os requisitos necessários para a publicação de dados ligados e para a construção de interfaces e aplicações – ou *Application Programming Interface* (API). Aborda

ainda as ontologias próprias para uso da comunidade de patrimônio cultural e apresenta um *framework* para conversão de dados de bases locais para o modelo RDF.

No mesmo sentido, Van Hooland e Verborgh (2014) tratam da modelagem, higienização, harmonização entre vocabulários controlados, enriquecimento e posterior publicação de dados de instituições *LAM*. Para cada um desses tópicos, apresentam estudos de caso relacionados a bibliotecas, arquivos ou museus.

O trabalho de Wood (*et al.*, 2014) por outro lado, constitui-se em um livro técnico voltado para desenvolvedores, em que são abordados aspectos específicos a respeito da construção de URIs, da modelagem de dados em RDF e suas formas de serialização, do protocolo SPARQL e de tipos de bases de dados que permitem a alocação de dados ligados. Neste caso, trata-se de uma obra geral que pode ser utilizada por outros domínios além do âmbito cultural.

A literatura científica nacional também apresenta suas contribuições para a temática. Santos Neto (*et al.*, 2013) apresenta um estudo de caso fictício de aplicação das ferramentas do movimento *Linked Open Data* para integração de dados de bibliotecas, arquivos e museus. A partir de registros distribuídos na *Web* sobre o escritor Machado de Assis e suas obras, Santos Neto (*et al.*, 2013) buscou ilustrar o processo de ligação de dados por meio do uso do modelo RDF e de outras ferramentas disponibilizadas pela comunidade de patrimônio cultural, como os dados de autoridades do consórcio *Virtual International Authority File* (VIAF) e os URIs disponibilizados pela *Library of Congress*, por exemplo.

No âmbito dos trabalhos de pós-graduação, a literatura nacional em Ciência da Informação registra três trabalhos sobre os princípios *Linked Data* e o movimento *Linked Open Data*. Arakaki (2016) trata do uso dos princípios *Linked Data* para integração de dados bibliográficos, realizando levantamento teórico a respeito da temática e apresentando iniciativas identificadas na literatura, em que se destacam os trabalhos da *Library of Congress*, da *Online Computer Library Center* (OCLC) e o projeto *Europeana*.

Sob outra perspectiva, Assumpção (2018) propõe um modelo para a ligação e publicação de dados de autoridade que envolve as políticas, os procedimentos e as tecnologias necessárias. Tais tecnologias são apresentadas e discutidas, bem como iniciativas identificadas, com destaque também para os serviços da *Library of Congress* e o consórcio VIAF. Este consórcio constitui, ainda, o objeto de estudo de Romanetto (2017), que trata do controle de autoridade por esta iniciativa e como o movimento *Linked Open Data* pode e tem contribuído para esse controle.

Finalmente, Costa (2017) apresenta sua contribuição para a temática, pois embora o seu trabalho se direcione aos campos dos Sistemas de Informação e da Gestão do Conhecimento<sup>17</sup>, tem como objetivo central a aplicação dos princípios *Linked Data* para a integração de dados de coleções museológicas. Nesse sentido, é um trabalho técnico e prático, em que efetivamente trabalhou-se com um conjunto de dados de coleções museológicas selecionados na Rede de Museus e Espaços de Ciências e Cultura da Universidade Federal de Minas Gerais (UFMG).

Todos estes trabalhos trazem contribuições a respeito das possibilidades de aplicação das ferramentas do movimento *Linked Open Data* para integração de dados de diferentes ambientes, respeitando-se as particularidades de cada um deles. Apontam também iniciativas e projetos já institucionalizados, bem como a apropriação de ferramentas e instrumentos por parte da comunidade de patrimônio cultural como, por exemplo, os disponibilizados pela *Library of Congress*.

A partir, então, do arcabouço apresentado por esses e outros autores, este capítulo apresenta movimentos e iniciativas que se inserem no escopo dos objetivos das Humanidades Digitais. O capítulo seguinte apresenta e discute aspectos dos padrões e modelos de dados desenvolvidos por essas e outras iniciativas para promover o acesso, uso e reuso de dados de patrimônio cultural.

A partir dos princípios *Linked Data* e do movimento *Linked Open Data*, profissionais e cientistas que trabalham com patrimônio cultural buscaram entender e aplicar os procedimentos de ligação e publicação de dados em suas instituições. Arakaki (2016) aponta os movimentos *Linked Open Data in Libraries, Archives and Museums* (LODLAM), *Linked Open Data in Galleries, Libraries, Archives and Museums* (LODGLAM) e *Open Galleries, Libraries, Archives and Museums* (Open GLAM).

O LODLAM refere-se ao movimento de instituições *LAM* em abordar, por meio de grupos de trabalho, as características do movimento *Linked Open Data* de forma a aplicá-las nesses ambientes (VOSS, 2012; ARAKAKI, 2016). O grupo se descreve como “[...] uma rede informal e sem fronteiras de entusiastas, técnicos, profissionais e outras pessoas interessadas

---

<sup>17</sup> Trabalho vinculado ao curso de Mestrado Profissional em Sistemas de Informação e Gestão do Conhecimento da Faculdade de Ciências Empresariais da Universidade FUMEC

em nosso trabalho com ‘*Linked open data*’ em relação às galerias, bibliotecas, arquivos e museus” (LODLAM.NET, 2010, *online*, tradução nossa<sup>18</sup>).

No mesmo sentido, LODGLAM é um movimento voltado ao patrimônio cultural, que inclui em sua nomenclatura as galerias de arte (ARAKAKI, 2016). Dentre os três movimentos, *Open GLAM* talvez seja o único institucionalmente formalizado, uma vez que é uma proposta da Comissão Europeia<sup>19</sup> coordenada pela organização sem fins lucrativos *Open Knowledge International*, e que tem como objetivo a disponibilização pública de conteúdos e de dados de galerias, bibliotecas, arquivos e museus (ARAKAKI, 2016). Em outras palavras, é um movimento voltado à promoção do livre acesso ao patrimônio cultural digital dessas instituições (OPEN KNOWLEDGE, [201-]b).

Os membros do movimento LODLAM elaboraram, durante um encontro do grupo em 2011, uma adaptação do *ranking* de Berners-Lee (2006) para a ligação e publicação de dados de patrimônio cultural. Esse *ranking*, denominado ‘*4-star classification scheme for linked open cultural metadata*’, trata especificamente dos aspectos de licenciamento dos metadados e *datasets* publicados, como explicado abaixo:

- Quatro estrelas: Dados e metadados publicados em domínio público. Nesta pontuação, os dados estão totalmente abertos, podendo ser ligados com outros dados (abertos ou de *datasets* restritos) e utilizados com fins comerciais ou não;
- Três estrelas: Dados e metadados publicados com atribuição de licença por meio de *linkbacks*. Neste caso, todos os usos dos dados e metadados publicados exigem a atribuição de indicação da fonte original desses dados. *Linkback* é um mecanismo simples e padronizado da plataforma *Web* que faz a ligação dos novos produtos/serviços com a fonte original dos dados;
- Duas estrelas: Dados e metadados publicados com outros meios de atribuição de licença. Nesta pontuação, o tipo de licença e o mecanismo utilizado para a atribuição são especificados pela fonte ou órgão publicador. Desse modo, a ligação de vários *datasets* entre si pode ser um processo complicado;

---

<sup>18</sup> “*LODLAM.net is an informal, borderless network of enthusiasts, technicians, professionals and any number of other people who are interested in or working with Linked Open Data pertaining to galleries, libraries, archives, and museums*”.

<sup>19</sup> Comissão interna da União Europeia, responsável por elaborar propostas legislativas para a Europa e implementar decisões do Parlamento Europeu (UNIÃO EUROPEIA, 2017)

- Uma estrela: Dados e metadados publicados com licença de compartilhamento comuns. Nesta pontuação, um determinado *dataset* só pode ser ligado a outros que contenham exatamente o mesmo tipo de licença, o que limita suas possibilidades de ligação (SMITH, 2011).

Ao contrário do *ranking* de Berners-Lee (2006), esta classificação não é cumulativa, de modo que as pontuações são independentes entre si. O grupo ainda optou por elencar primeiro a maior classificação, que diz respeito ao acesso, uso e reuso de dados e metadados sem restrições, diminuindo as demais pontuações conforme essas restrições se tornam maiores e dificultam o compartilhamento de dados e metadados. Percebe-se que com este *ranking* o movimento LODLAM tem como objetivo potencializar o compartilhamento e o uso irrestrito de dados de patrimônio cultural publicados no escopo do movimento *Linked Open Data*.

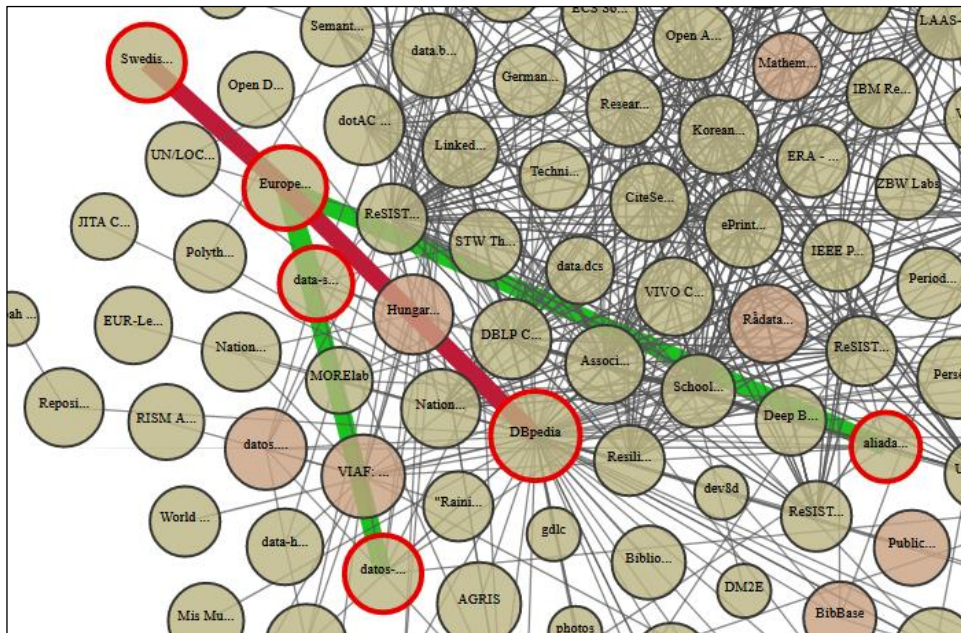
Em âmbito institucional, destacam-se também iniciativas internacionais. O projeto Europeana, por exemplo, é uma iniciativa da Fundação para a Biblioteca Digital Europeia<sup>20</sup>, em parceria com o Parlamento Europeu e sediado na Biblioteca Nacional dos Países Baixos (Holanda), que agrega e disponibiliza dados do patrimônio cultural de todos os estados-membros da União Europeia, datados do período pré-histórico até os atuais, em 29 idiomas (WINER; ROCHA, 2013).

Coneglian e Santarem Segundo (2017) afirmam o protagonismo desse projeto no cenário mundial, pois este contabiliza atualmente mais de 50 milhões de objetos culturais de instituições *LAM*. Estes objetos podem ser acessados tanto por meio de um portal como por meio de um SPARQL API. O *dataset* Europeana está inserido no projeto *LOD-Cloud*, conforme a figura 7.

---

<sup>20</sup> Fundação da Comissão Europeia para a criação da “Biblioteca Europeia”, antecessora da Europeana (EUROPEANA, 2017b).

**Figura 7 - Dataset Europeana no projeto LOD-Cloud**



Fonte: McCrae (et al., 2018).

Utilizando o modo interativo da visualização de *datasets* do projeto *LOD-Cloud* é possível identificar o conjunto de dados Europeana, no canto superior esquerdo da figura 7, bem como os *datasets* que se relacionam a ele, por meio de *links* de entrada (em vermelho) e de saída (verde). Percebe-se, por exemplo, que a Europeana fornece dados ao *dataset* DBpedia (Wikipedia). O *dataset* Europeana faz parte do domínio “publicações”.

A abrangência do projeto e a complexidade em reunir dados de diferentes países levou à criação de um modelo de dados próprio, o *Europeana Data Model* (EDM), que é utilizado em conjunto a outros padrões e vocabulários do consórcio W3C, de modo que o projeto “[...] adota um *framework* aberto, com domínios cruzados e baseado em *Web Semântica*, que acomoda a riqueza e o alcance de padrões de comunidades específicas” (EUROPEANA, 2013, p.5, tradução nossa<sup>21</sup>). Este modelo será apresentado e discutido no próximo capítulo.

É importante ressaltar, finalmente, que existem objetos agregados pelo projeto que fazem referências ao patrimônio cultural brasileiro de diferentes períodos históricos, fato que pode constituir ponto fundamental para iniciativas em âmbito nacional de criação de projetos semelhantes (WINER; ROCHA, 2013). Tais iniciativas poderiam, ainda, firmar parcerias com o projeto europeu.

<sup>21</sup> “EDM [...] adopts an open, cross-domain Semantic Web-based framework that can accommodate the range and richness of particular community standards”

A partir do projeto Europeana apontam-se outras iniciativas dos países europeus que se reportam a ele. Cagnazzo (2017), bem como Winer e Rocha (2013), elencam os países que mais contribuem com conteúdo para esse projeto, que são: Alemanha, França, Holanda, Suécia, Espanha, Inglaterra, Noruega, Polônia e Itália.

Nesse sentido, Cagnazzo (2017) descreve em sua pesquisa de dissertação os resultados de entrevistas que fez com alguns centros informacionais incorporados pela Europeana, que seguem:

- *British Library*: com as políticas favoráveis para tornar os dados públicos, a *British Library* encontrou nos princípios *Linked Data* a oportunidade de realizar a disponibilização dos seus dados. Além disso, utilizou-se desses princípios para atualizar as tecnologias de disponibilização de suas coleções em relação aos registros bibliográficos e de autoridade. Com uso do modelo EDM, a *British Library* disponibiliza uma seleção dos metadados, nomeada como ‘*Researcher Format*’, que facilita ao usuário a visualização das informações;
- *National Library of Finland*: já disponibiliza dados de autoridade, mas futuramente seus dados bibliográficos serão expostos. A biblioteca adota o *Schema.org* como modelo de dados para publicação;
- *National Library of Germany*: esta biblioteca iniciou seus esforços pela modelagem dos dados de autoridade por meio do uso de vocabulários comuns, das orientações do *Resource Description and Access* (RDA) e de elementos dos *namespaces Simple Knowledge Organization System* (SKOS) e DCTERMS. Ao mesmo tempo, o projeto *Gemeinsame Normdatei* (GND), ou *Integrated Authority File*, tem como objetivo integrar esses dados em uma única base de dados, que até então estavam alocados em bases separadas para pessoas, entidades coletivas, termos de assuntos e trabalhos criativos - especialmente música;
- *National Library of Latvia* e *National Library of Luxembourg*: nessas bibliotecas, os dados de autoridade são publicados como *Linked Data* por meio do envio dos registros MARC para o VIAF e também para a Europeana;
- *National Library of the Netherlands*: desenvolveu um modelo de dados baseado no *Bibliographic Framework Initiative* (BIBFRAME), da *Library of Congress* (LC). Este modelo é fundamentado no *Schema.org* para que motores de busca, como o *Google*, encontrem os registros de suas coleções;

- *National Library of France*: também desenvolveu um modelo de dados flexível, com uso do modelo EDM e dos padrões W3C, isto é, do modelo RDF e da linguagem XML, para a descrição dos dados;
- *Open Knowledge Greece* e *National Library of Greece*: iniciaram projetos sobre publicação *Linked Data*, estudando modos de transformar seus catálogos de autoridade em triplas RDF, porém não disponibilizaram nenhum *dataset* até o momento;
- *National Library of Spain*, *National Library of Portugal*, *National Library of Wales*, *National Library of Czech Republic*, *National Central Library of Florence*, *National Library of Liechtenstein* e *National Library of Bulgaria*: essas instituições iniciaram pesquisas sobre enriquecimento de dados de coleções de patrimônio cultural, mas têm encontrado dificuldades em várias questões como, por exemplo, infraestrutura, condições do acervo e de recursos humanos, que por sua vez envolvem a falta de funcionários capacitados e do interesse desses em aprimorar o acesso às coleções dessas instituições (CAGNAZZO, 2017).

Di Giorgio (2014) trata ainda de outro projeto vinculado à Europeia, o Projeto *Culturaitalia* (Portal da Cultura Italiana<sup>22</sup>). Esse projeto agrega nacionalmente os dados de patrimônio cultural de instituições italianas de âmbito nacional, regional e local, e está diretamente vinculado ao Ministério do Patrimônio Cultural, Atividades e Turismo (MiBACT) da Itália (DI GIORGIO, 2014).

Destacam-se também projetos sediados em outros continentes. O projeto *Digital New Zealand* (*DigitalNZ*) é uma iniciativa do governo neozelandês, liderado pela Biblioteca Nacional da Nova Zelândia, de disponibilização de dados e metadados de patrimônio cultural (DIGITALNZ, 2017). Com pouco menos de 31 milhões de itens, o *DigitalNZ* assemelha-se ao projeto Europeia por agregar e permitir a busca e o compartilhamento de registros de metadados. Este projeto destaca-se pelo incentivo à digitalização e inclusão de conteúdos pela própria comunidade que se utiliza do portal e do *DigitalNZ* API.

Dentre suas fontes de dados de patrimônio cultural estão o próprio projeto Europeia, o projeto *Digital Public Library of America* (DPLA), bem como outras instituições culturais alocadas na Nova Zelândia, como o arquivo *Hutt City Council*, a biblioteca Alexander Turnbull e a Universidade Lincoln (DIGITALNZ, [201-]b).

---

<sup>22</sup>.[dati.culturaitalia.it](http://dati.culturaitalia.it)



No continente asiático destaca-se a união das instituições da Coreia<sup>23</sup> em torno do projeto de pesquisa ‘*K-Culture Time Machine*’, cujos objetivos são o de relacionar diferentes bases de dados de patrimônio cultural por meio de um esquema de metadados, denominado *Korean Cultural Heritage Data Model* (KCHDM), e o desenvolvimento de uma ontologia para um serviço de informações do patrimônio do país (KIM *et al.*, 2015).

No continente americano, por outro lado, destaca-se a supracitada iniciativa DPLA<sup>24</sup>. Esta é uma biblioteca digital que agrega recursos de instituições *LAM* dos Estados Unidos, também de forma semelhante ao projeto Europeana, e seus dados publicados visam o reuso e o desenvolvimento de aplicativos (THORSEN; PATUELLI, 2016). O acesso aos seus dados também ocorre por meio de um API. Assim como o projeto europeu, o acervo da DPLA contém referências ao patrimônio cultural brasileiro, de modo que pode se constituir também como um projeto parceiro em iniciativas nacionais.

Finalmente, atenta-se para o projeto *Google Arts & Culture*, que tem como objetivo “disponibilizar *online* o patrimônio cultural mundial” (GOOGLE CULTURAL INSTITUTE, 2018). O projeto foi elaborado pelo *Google Cultural Institute*, fundação sem fins lucrativos criada em 2011, e oferece diferentes modos de acesso ao patrimônio cultural digital agregado pelo projeto, destacando-se o uso de realidade virtual para ‘visitação’ a museus e a visualização de imagens de obras de arte, objetos e documentos em alta qualidade. Dentre as coleções de patrimônio cultural disponibilizadas por esse projeto encontram-se coleções de instituições brasileiras como, por exemplo, a Pinacoteca do Estado de São Paulo e o Museu de Arte Moderna do Rio de Janeiro.

Se no continente europeu e na América do Norte existem inúmeras iniciativas de publicação e ligação de dados de patrimônio cultural, a realidade brasileira parece não acompanhar a tendência. O Ministério da Cultura (MinC) anunciou seu plano de abertura de dados, que vai disponibilizar, em etapas, dados administrativos e burocráticos de suas secretarias, órgãos vinculados<sup>25</sup>, do Sistema de Apoio às Leis de Incentivo à Cultura (Salic), do Mapa da Cultura, do Sistema Nacional de Informações e Indicadores Culturais (SNIIC), do Vale-Cultura, da Rede Cultura Viva, do Sistema Nacional de Cultura (SNC) e dos Centros de Artes e Esportes Unificados (CEUs) (BRASIL, 2017b). A única menção, porém, ao

---

<sup>23</sup> Denominação usada pelos autores, referindo-se à República da Coreia, ou Coreia do Sul.

<sup>24</sup> dp.la

<sup>25</sup> Instituto Brasileiro de Museus (Ibram), Fundação Nacional de Artes (Funarte), Agência Nacional do Cinema (Ancine), Fundação Casa de Rui Barbosa (FCRB), Fundação Biblioteca Nacional (FBN), IPHAN e Fundação Cultural Palmares.

patrimônio cultural nacional é relativa ao Cadastro de Bens Culturais Musealizados Desaparecidos (CBMD), que terá sua listagem publicada e periodicamente atualizada no portal de dados abertos do MinC<sup>26</sup> (BRASIL, 2017a).

Ao contrário da Biblioteca Nacional e do IPHAN, o Arquivo Nacional está vinculado ao Ministério da Justiça e Segurança Pública (MJSP). O órgão também anunciou seu plano de abertura de dados administrativos, mas não faz menção à disponibilização de dados de seu acervo histórico, embora possua como uma diretriz de planejamento estratégico a digitalização de documentos, com o objetivo de preservar a memória nacional (BRASIL, 2017c).

Esses movimentos nacionais relacionam-se aos princípios *Open Data* aplicados aos dados governamentais, tendência global de incentivo à transparência pública que impulsionou, no Brasil, a criação da Lei de Acesso à Informação<sup>27</sup> (SANTAREM SEGUNDO, 2013). Porém, não foram identificados indícios de formação de grupos de trabalho, ou de iniciativas relativas à aplicação dos princípios do movimento *Linked Open Data* em dados de patrimônio cultural nos *sites* da Biblioteca Nacional<sup>28</sup>, do IPHAN<sup>29</sup>, ou do Arquivo Nacional<sup>30</sup>.

O IBRAM, por outro lado, tem trabalhado com o próprio MinC e com a Universidade Federal de Goiás no Projeto Tainacan, que tem como objetivo a integração e gestão digital interoperável de acervos de bibliotecas, arquivos, museus e também de cinematecas, por meio do desenvolvimento de uma plataforma homônima (IBRAM, [201-]). Ainda em fase inicial, porém, o projeto possui pouca documentação e não faz menção clara aos princípios *Linked Data* ou ao movimento *Linked Open Data*.

Diante do número de iniciativas internacionais, nota-se que as instituições de guarda de patrimônio cultural nacionais possuem, então, uma lacuna no que diz respeito à tendência global de disponibilização de dados abertos e ligados desses objetos.

Nesse sentido, o próximo capítulo apresenta e discute os modelos de dados, os padrões de metadados e outros instrumentos desenvolvidos pela comunidade de patrimônio cultural para a descrição *Linked Open Data* e a disponibilização de dados de coleções de patrimônio cultural que podem ser utilizadas por instituições *LAM* que queiram iniciar esses processos.

---

<sup>26</sup> dados.cultura.gov.br

<sup>27</sup> Lei nº 12.527, de 18 de novembro de 2011 (BRASIL, 2011).

<sup>28</sup> www.bn.gov.br

<sup>29</sup> www.iphan.gov.br

<sup>30</sup> www.arquivonacional.gov.br

## 5 MODELOS DE DADOS E PADRÕES DE METADADOS PARA DESCRIÇÃO *LINKED OPEN DATA* DE COLEÇÕES DE PATRIMÔNIO CULTURAL

Uma vez elencados os aspectos de representação informacional no ambiente digital, por meio do uso de padrões de metadados, e as características de descrição de recursos em relação ao movimento *Linked Open Data*, é possível identificar esforços da comunidade de patrimônio cultural em harmonizar seus procedimentos e instrumentos em relação à descrição e disponibilização de dados de suas coleções.

Assim, este capítulo retoma as questões sobre os padrões de metadados para a descrição de coleções de patrimônio cultural, apresentados no capítulo 3. Nesse contexto, são apresentadas as soluções desenvolvidas pelas instituições de patrimônio cultural para uso desses padrões na descrição de suas coleções no escopo do movimento *Linked Open Data*.

Inicialmente, porém, expõem-se considerações sobre os modelos de dados desenvolvidos pelas iniciativas Europeia, DPLA e *DigitalNZ* sob o ponto de vista do escopo de atuação dessas iniciativas, das tecnologias e ferramentas utilizadas e da própria configuração dos seus modelos de dados. Essas características foram levantadas na literatura documental pertinente a cada iniciativa e comparadas com o objetivo de identificar suas semelhanças e diferenças.

### 5.1 Modelos de dados

Considerando a relevância dos projetos Europeia, DPLA e *DigitalNZ* na disponibilização de dados estruturados de coleções de patrimônio cultural, buscam-se identificar e discutir os modelos de dados adotados por essas iniciativas para o mapeamento e descrição de suas coleções. A seleção desses projetos guiou-se pela identificação, na literatura documental pertinente a cada um deles, do uso de APIs para a promoção do acesso aos dados estruturados de acordo com os princípios *Linked Data*.

O projeto DPLA define uma API como

[...] um conjunto de métodos para a comunicação com um sistema de *software* (frequentemente complexo) e a obtenção de respostas de forma que um computador (e, com alguma prática, um ser humano) possa entender. Em essência, uma API define a ‘linguagem’ que um sistema fala. Como uma linguagem, ela tem seu próprio vocabulário com termos que têm significados especiais (por exemplo, nomes e rótulos de propriedade), gramática (como esses nomes e rótulos de propriedade são organizados - coletivamente, seu

esquema) e sintaxe (ou seja, como a própria informação está organizada). Como uma linguagem, você pode usá-la para **fazer perguntas e entender respostas** (DPLA, [201-]a, *online*, grifo e tradução nossa<sup>31</sup>).

Percebe-se que esses projetos utilizam as APIs como ferramentas para a busca e a recuperação de dados. Nesse sentido, apresentam-se também breves considerações sobre as soluções de consulta e de acesso aos dados desenvolvidas por cada um desses três projetos.

O projeto Europeana desenvolveu o modelo de dados EDM. Seu objetivo é ser um *framework* para coleta, conexão e enriquecimento de metadados (EUROPEANA, [201-]). Este modelo substitui o *Europeana Semantic Elements* (ESE), primeiro modelo de dados desenvolvido e que, ao contrário do EDM, não considerava as particularidades bibliográficas, museológicas ou arquivistas dos objetos e dos padrões de metadados que os descrevem (EUROPEANA, 2013).

Este projeto funciona como um agregador de dados de diferentes instituições, que muitas vezes trabalham com padrões e procedimentos de descrição heterogêneos. Nesse sentido, o EDM reutiliza elementos de *namespaces* como RDF, RDF-S (ontologia em RDF), *Dublin Core*, *Creative Commons* (CC), *OAI Object Reuse and Exchange* (OAI-ORE), dentre outros, em consonância com seus elementos próprios para a descrição de recursos no escopo do projeto (EUROPEANA, 2017b).

Dentre esses *namespaces*, destacam-se os relativos ao *Creative Commons* e à solução OAI-ORE. Esta última “[...] define padrões para a descrição e troca de agregações de recursos *Web*” (OPEN ARCHIVES INITIATIVE ([2008?], *online*, tradução nossa<sup>32</sup>). Assim, sua utilização trata de um aspecto de especial importância para o projeto Europeana: a agregação de recursos digitais relativos a determinado objeto físico e, conseqüentemente, com as diferentes informações sobre ele distribuídas entre as instituições participantes.

Em relação aos direitos autorais e à permissão de uso de dados, o projeto optou, então, pelo uso do *Creative Commons*, que trata especificamente de aspectos de licenciamento jurídicos gratuitos que permitem diferentes modos de compartilhamento de conteúdo

---

<sup>31</sup> *At its most basic, an API, or Application Programming Interface, is made up of a set of defined methods that someone can use to communicate with a (frequently complex) software system, and get back responses in a way that a computer (and, with some practice, a human) can understand.*

*In essence, an API defines the ‘language’ a system speaks. Like a language, it has its own vocabulary with terms that have special meanings (e.g., property names and labels), grammar (how those property names and labels are arranged—collectively, its schema), and syntax (i.e., how the information itself is arranged). Like a language, you can use it to ask questions and understand responses*

<sup>32</sup> *Open Archives Initiative Object Reuse and Exchange (OAI-ORE) defines standards for the description and exchange of aggregations of Web resources.*

(CREATIVE COMMONS, [20--]). Sobre os aspectos específicos de licenciamento de dados abertos, destacam-se ainda a possibilidade de uso do *Open Data Commons*, que promove ferramentas legais para publicação e uso de dados (OPEN KNOWLEDGE, [201-]a).

Os elementos próprios do modelo EDM compõem o *namespace edm* e, em conjunto com os demais, modelam os dados de acordo com os requisitos básicos elencados abaixo:

- Distinção entre os recursos e suas representações digitais;
- Distinção entre os recursos e os metadados que o descrevem;
- Possibilidade de um mesmo objeto ter múltiplos registros, mesmo que estes apresentem descrições conflitantes entre si;
- Suporte a recursos que são compostos por outros recursos;
- Compatibilidade com diferentes níveis de abstração de descrições;
- Padronização do formato de metadados, podendo ser especializado;
- Suporte para recursos contextuais, como vocabulários controlados (CONEGLIAN; SANTARÉM SEGUNDO, 2017; EUROPEANA, 2013).

O primeiro e o segundo requisitos remetem às considerações de Glushko (2013) sobre a organização de objetos e de suas informações como procedimentos separados. Desse modo, o projeto trabalha com uma classe específica que promove a agregação de descrições ricas e, muitas vezes, redundantes ou conflitantes de recursos, de acordo com o terceiro, quarto e quinto requisitos, mas que respeitam as características individuais do recurso em si e de suas representações digitais.

O sexto requisito trata da formatação dos dados. O EDM faz uso linguagem XML e desenvolveu o *EDM XML Schema* (EUROPEANA, 2014), que utilizado em conjunto com o modelo RDF (RDF/XML) permite a publicação do *dataset* Europeia no projeto *LOD-Cloud* (CONEGLIAN; SANTARÉM SEGUNDO, 2017).

Finalmente, o sétimo requisito diz respeito ao uso dos padrões de valores de dados, ou vocabulários controlados. O modelo permite, então, que esses vocabulários sejam inseridos e utilizados para a descrição dos recursos.

A partir desses requisitos, assinalam-se os princípios de *design* do modelo EDM:

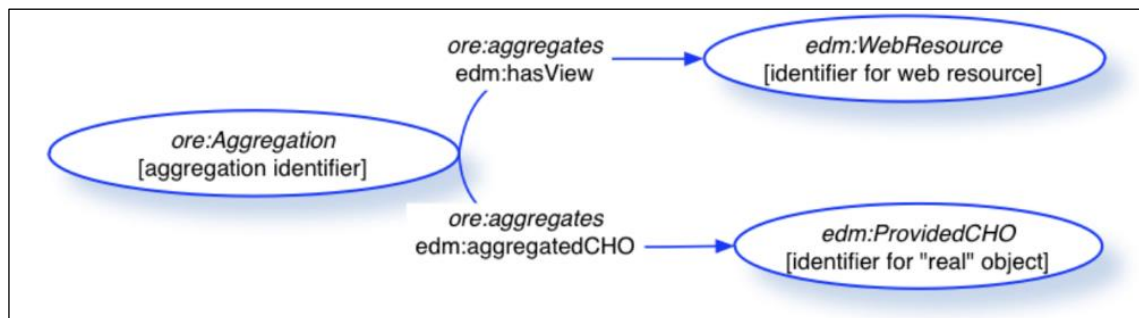
- Integração de dados em ambiente aberto, uma vez que não é possível antecipar todas as contribuições de dados;

- Rica funcionalidade, com a possibilidade de extensões e
- Reutilização de modelos e padrões existentes (EUROPEANA, 2013)

Estes princípios evidenciam o aspecto colaborativo entre o projeto e as diferentes instituições provedoras de dados. O projeto Europeana, portanto, segue as recomendações do consórcio W3C no que diz respeito ao aproveitamento de padrões e ferramentas consolidados pela comunidade e sobre manter um ambiente aberto e colaborativo de trabalho.

Nesse sentido, os elementos do EDM estão divididos em três classes gerais “obrigatórias” de recursos, cinco classes contextuais e diversas propriedades. As classes gerais tratam especificamente do objeto/recurso em si (*edm:ProvidedCHO*), suas representações digitais (*edm:WebResource*) e uma agregação (*ore:Aggregation*) que representa a união lógica entre as duas primeiras (EUROPEANA, 2013). A figura 8 apresenta essas três classes, com as propriedades básicas que as relacionam.

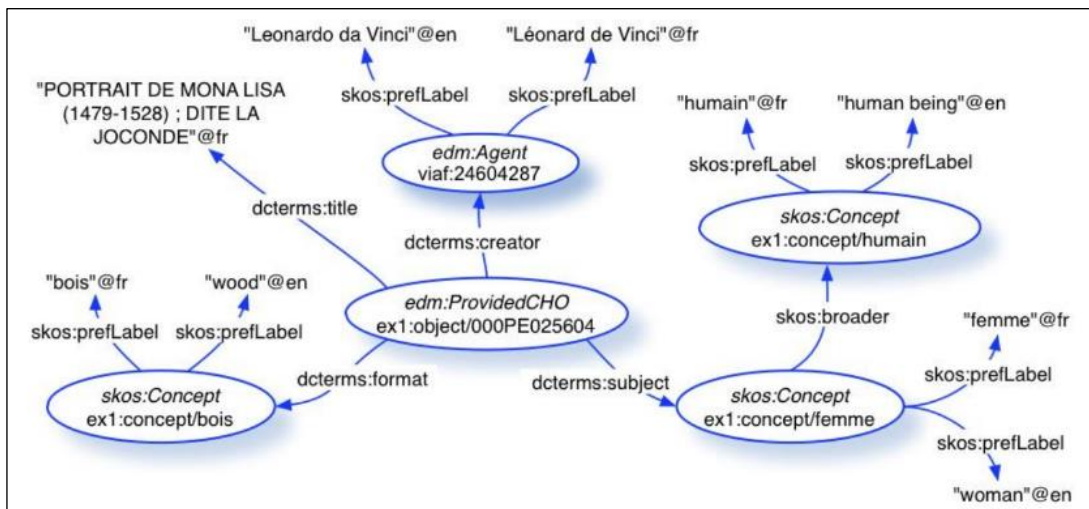
**Figura 8** - Classes principais do modelo EDM



Fonte: Europeana (2013)

As classes contextuais *edm:Agent*, *edm:Place*, *edm:TimeSpan*, *skos:Concept* e *cc:License*, como a própria denominação indica, contextualizam o recurso a respeito do “quem”, “onde”, “quando”, “o quê” e “acesso e uso”, respectivamente. A figura 9, apresenta um exemplo de uso das classes *edm:Agent* e *skos:Concept* para a descrição da obra “Mona Lisa”.

**Figura 9** - Exemplo de uso das classes contextuais do EDM



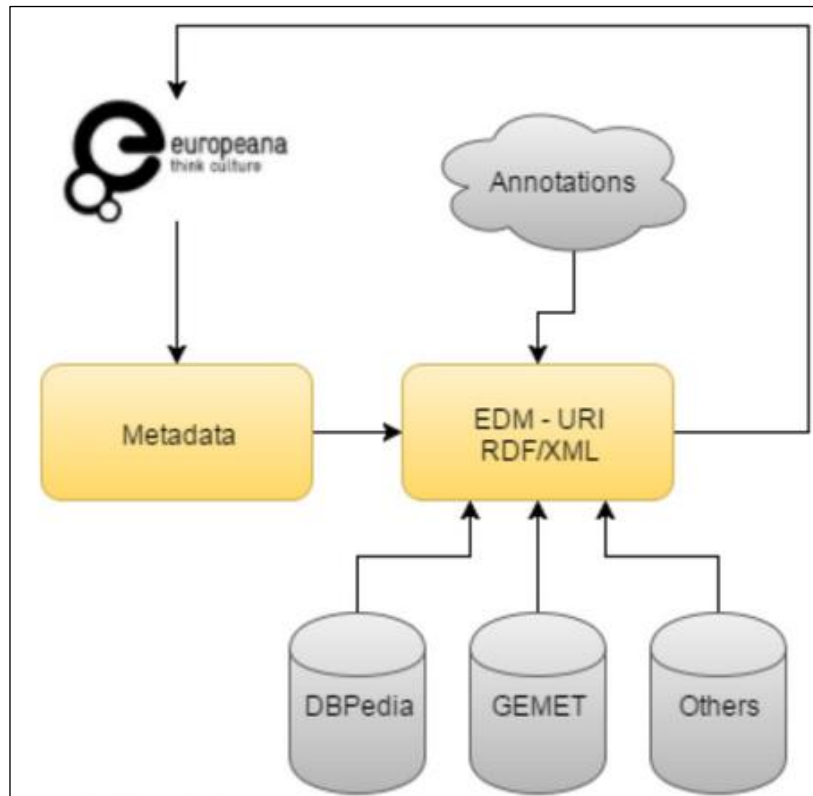
Fonte: Europeana (2013)

Neste exemplo, a partir do objeto real (*edm:ProvidedCHO*) indicam-se suas propriedades ‘título’ (*dcterms:title*), ‘criador’ (*dcterms:creator*) - por meio do uso da classe *edm:Agent* - ‘assunto’ (*dcterms:subject*) e ‘forma física’ (*dcterms:format*) - por meio do uso da classe *skos:Concept*.

A partir dessa modelagem, o projeto Europeana (2017a) recomenda que a descrição dos recursos seja a mais específica possível, esclarecendo ainda que o EDM permite representações orientadas a objetos ou a eventos, em conjunto ou isoladamente, além de outros modos de descrição.

Desse modo, Coneglian e Santarem Segundo (2017, p.98) consideram que o EDM “[...] apresenta-se como um dos principais modelos de informação cultural de acesso público e semanticamente organizado disponíveis na *Web*”, uma vez que o projeto é eficiente em agregar e publicar dados de forma estruturada e com relações com outros *datasets*, como demonstrado na figura 10.

**Figura 10** - Esquema de conversão de metadados para o modelo EDM



Fonte: Coneglian e Santarem Segundo (2017)

Coneglian e Santarem Segundo (2017) explicam que os metadados coletados nas fontes primárias estão alocados na base de dados do projeto e são convertidos para o modelo EDM por meio da aplicação de folhas de estilo, que atribuem URIs e fazem as representações no modelo RDF e no formato RDF/XML. Posteriormente, têm-se o enriquecimento de dados, que incluem anotações sobre os recursos feitas pelo escritório da Europeana e posterior relacionamento dos dados com outros *datasets*, como o *DBpedia* e o *General Multilingual Environmental Thesaurus* (GEMET), indicados na figura 10.

Além de publicar seu *dataset* no projeto *LOD-Cloud*, o projeto Europeana fornece acesso aos seus dados por meio de diferentes APIs. Para a recuperação de dados estruturados, utiliza-se a *SPARQL* API (EUROPEANA, 2018), que como a própria denominação indica, faz uso do protocolo *SPARQL* para a configuração das consultas aos dados.

O reconhecimento da eficiência do EDM evidencia-se na exploração desse modelo por outras iniciativas como, por exemplo, o projeto DPLA. Este desenvolveu um modelo de dados em RDF denominado *DPLA Metadata Application Profile* (DPLA - MAP), com o objetivo de aplicá-lo em conjunto ao DPLA API (DPLA, 2017b). Atualmente, o modelo está na sua quinta versão, lançada em 2017.

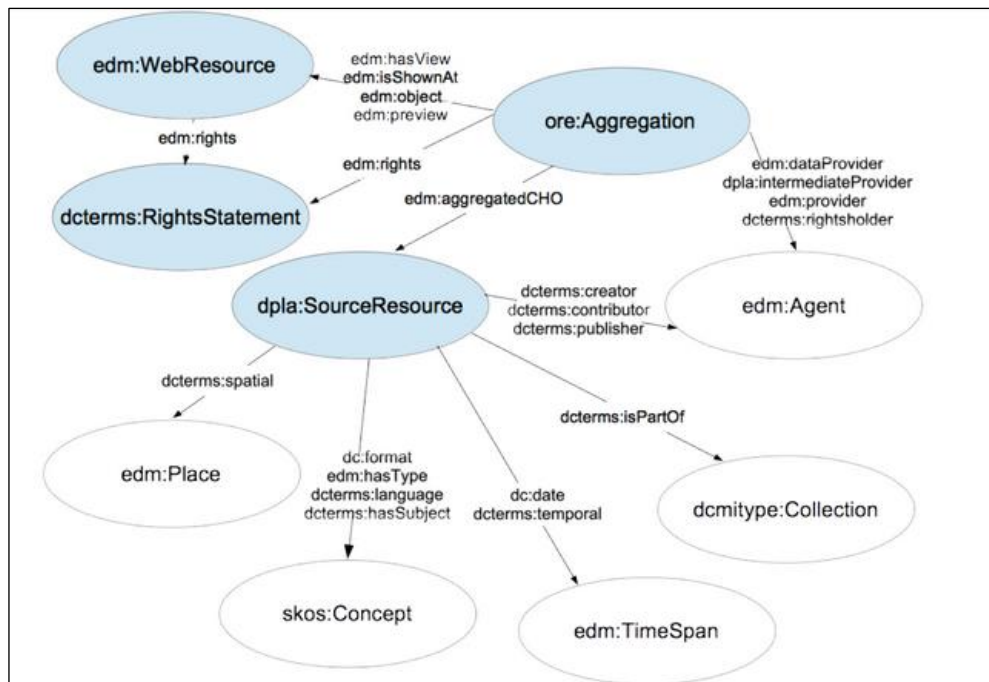


Tal modelo possibilita a integração dos metadados criados, agregados ou compartilhados por outras instituições em vários padrões, ao mesmo tempo em que se constitui em um modelo semântico e abstrato que descreve os relacionamentos entre os dados de determinado recurso (DPLA, 2017a). O DPLA MAP, então, fornece acesso aos dados de recursos alocados em diferentes instituições ao mesmo tempo em que aponta para os relacionamentos entre esses recursos.

Nessa perspectiva, o projeto norte-americano esclarece que seu escopo inclui apenas o *harvesting* de metadados, encaminhando os usuários à fonte de origem dos recursos, e que, por isso, reutiliza apenas parte do modelo EDM (DPLA, 2017a).

O DPLA MAP também reutiliza elementos dos *namespaces Dublin Core*, SKOS e ORE, além do próprio EDM, e sua estrutura envolve quatro classes principais e cinco classes contextuais, como demonstrado na figura 11.

**Figura 11** - Classes principais e contextuais do modelo DPLA MAP



Fonte: DPLA (2017a)

Percebe-se que o DPLA MAP se utiliza da classe *ore:aggregation* para empacotar informações sobre os relacionamentos das demais classes principais - na figura 10, estas classes estão coloridas na cor azul. Os relacionamentos entre as classes principais fornecem dados sobre o armazenamento de versões digitais do recurso (*edm:WebResource*) e sobre aspectos de licenciamento e compartilhamento (*dcterms:RightsStatement*) (DPLA, 2017a).

Percebe-se o contraponto com o projeto Europeia, que considera os aspectos de licenciamento como informações contextuais.

Destaca-se também o uso da classe principal *dpla:SourceResource*. Esta classe agrupa, por meio do relacionamento com as classes contextuais *edm:Agent*, *dcmitype:Collection*, *skos:Concept*, *edm:Place* e *edm:TimeSpan*, as propriedades que fornecem os metadados descritivos dos recursos (DPLA, 2017a). Nesse contexto, pode-se considerar que esta classe também funciona como uma agregação, no caso, de metadados descritivos contextuais. Tais metadados incluem informações sobre o pertencimento, ou não, de um recurso a uma determinada coleção (*dcmitype:Collection*).

Por ser um modelo abstrato, o DPLA MAP pode ser expresso em qualquer linguagem padronizada. No mesmo sentido, o projeto ainda afirma que as características de estruturação de seu modelo permitem a compatibilidade com qualquer padrão de metadados, inclusive VRA Core, CDWA e CIDOC *Conceptual Reference Model* (CIDOC-CRM) (DPLA, 2017a).

Originalmente, os arquivos de dados estão no formato JSON e, para o uso via API, JSON-LD, que fornece a estruturação dos dados como dados ligados. Nesse sentido, o acesso a esses dados acontece por meio do DPLA API. Este permite a recuperação de itens ou de coleções de itens que satisfazem aos requisitos de busca solicitados (DPLA, [201-]a).

Ressalta-se que o projeto ainda permite o *download* ‘em massa’, ou *bulk download*, de todo o seu repositório de dados, também no formato JSON (DPLA, [201-]b). Assim, embora esta iniciativa não disponibilize seu *dataset* no escopo do projeto *LOD-Cloud*, como no caso da Europeia, evidencia-se que o projeto DPLA preocupa-se com a disponibilização de seus dados para uso da comunidade.

Finalmente, o projeto *DigitalNZ*, assim como o DPLA, tem como foco a agregação de de metadados. Seu sistema de *harvesting*, denominado *Supplejack*, busca, agrega e compartilha registros de metadados em diferentes formatos, também por meio de um API (DIGITALNZ, [2017]b). Esta ferramenta foi criada por um grupo de profissionais da Biblioteca Nacional da Nova Zelândia e do Departamento de Assuntos Internos do governo para, inicialmente, agregar coleções de bibliotecas, arquivos, museus, comunidades (dados disponibilizados por usuários), governos e radiodifusores do país, mas transformou-se em uma plataforma de uso geral que consegue agregar milhões de dados, provenientes de diferentes fontes (SUPPLEJACK, [2013]). Desse modo, constitui-se em uma ferramenta que pode ser livremente utilizada em outros projetos que envolvam instituições LAM.

Dentre as funcionalidades do *Supplejack* estão: 1) a possibilidade de definição customizada de um esquema de dados para mapeamento de dados de entrada, 2) o suporte

para manipulação, validação, *namespacing*, transformação e enriquecimento de dados, e 3) o suporte de vários formatos, protocolos, modelos e padrões de dados, dentre eles XML, OAI-PMH, HTML, RDF e JSON (SUPPLEJACK, [2013]). Trata-se, então, de uma ferramenta que permite a execução de diversos processos em um único ambiente e de forma personalizada.

A partir, então, da possibilidade de personalização do esquema de dados para o mapeamento dos dados de entrada, o projeto *DigitalNZ* elaborou um esquema de metadados específicos, cujo *namespace* é “dnz”, e que é utilizado em conjunto ao padrão *Dublin Core* para o mapeamento e *harvesting* de dados (DIGITALNZ, [201-]c). Cada um dos campos de metadados, independentemente do *namespace*, possui três características primordiais:

- *status* do requisito: determina se o campo é obrigatório, recomendável ou opcional;
- disponibilidade: determina se o campo é recuperável por máquina (via API) e
- valoração: especifica se o campo é uni ou multivalorado (DIGITALNZ, [201-]c).

Considerando que a recuperação e o acesso aos dados coletados se dão por meio de um API, os aspectos de disponibilidade possuem relevância para o processo de mapeamento dos dados a serem coletados. A valoração, por outro lado, auxilia na manutenção da consistência dos dados, uma vez que especifica quais campos podem ter apenas um valor – número identificador único, por exemplo - e quais podem ter mais de um valor, como itens que foram criados por mais de um autor, por exemplo.

Um aspecto diferencial do sistema *Supplejack* é a proposta de inclusão, ainda em fase experimental, de agregações relativas aos dados de autoridades e de lugares, denominadas *Concepts*, no seu modelo de dados (SUPPLEJACK, [2013], DIGITALNZ, [2017]b). Essas agregações reuniriam todos os dados, ou informações, relativos à determinada autoridade ou local em uma única entidade, assim como ocorre com as agregações de recursos pelos projetos Europeana e DPLA. Do ponto de vista do controle de autoridades e de consistência de informações descritivas, tal proposta diferencia o projeto *DigitalNZ* em relação aos outros. Ressalta-se, porém, que dentre toda a documentação apresentada por esse projeto esta é a única menção a um modelo de dados.

Em relação ao acesso aos dados pelos usuários e por outras instituições, o projeto disponibiliza o *DigitalNZ* API. Este se encontra em sua terceira versão e permite o acesso aos dados das coleções de três modos: pela recuperação de ‘sets’ (*datasets* formados pelas preferências dos usuários) no formato JSON, pela recuperação do conjunto de metadados

relativos a um determinado item e pela recuperação de itens que atendam a um determinado requisito de busca (DIGITALNZ, [201-]a).

Apesar de todos esses esforços, o projeto esclarece que ainda está trabalhando na sua estrutura *Linked Open Data*, por meio da construção de uma estrutura sustentável para os conteúdos incluídos pelos próprios usuários (DIGITALNZ, 2017). Porém, diferentemente dos projetos Europeana e DPLA, as documentações do sistema *Supplejack* e do *DigitalNZ* API são escassas e voltadas para profissionais da computação, dificultando os estudos sobre a estrutura de seu esquema de mapeamento de dados, ou mesmo sobre a modelagem de dados.

O quadro 6 sintetiza, a partir da análise da literatura, as características levantadas em relação ao escopo, às tecnologias e aos modos de descrição e acesso aos dados adotados pelos projetos Europeana, DPLA e *DigitalNZ*.

**Quadro 6** - Comparação entre os projetos Europeana, DPLA e DigitalNZ

		<b>Europeana</b>	<b>DPLA</b>	<b>DigitalNZ</b>
<b>Escopo</b>		Agregação de conteúdos	<i>Harvesting</i> de metadados	<i>Harvesting</i> de metadados
<b>Modelo de dados</b>		EDM	DPLA MAP	Esquema de mapeamento de dados do sistema <i>Supplejack</i>
<b>Tecnologias</b>	Ferramentas	URI, RDF, <i>SPARQL</i>	RDF	RDF
	Estrutura	RDF/XML	JSON JSON-LD	JSON
	<i>namespaces</i>	<i>rdf, rdf-s, Dublin Core, cc, ore, skos, owl, etc</i>	<i>edm, Dublin Core, skos, ore</i>	<i>dnz e Dublin Core</i>
<b>Descrição</b>	Classes	Principais: <i>edm:WebResource</i> <i>edm:ProvidedCHO</i> <i>ore:Aggregation</i> Contextuais: <i>edm:Agent,</i> <i>edm:Place</i> <i>edm:TimeSpan</i> <i>skos:Concept</i> <i>cc:Licence</i>	Principais: <i>edm:WebResource,</i> <i>ore:Aggregation</i> <i>dcterms:Rightstatement</i> <i>dpla:SourceResource</i> Contextuais: <i>edm:Agent</i> <i>edm:Place</i> <i>edm:TimeSpan</i> <i>skos:Concept</i> <i>dcmitype:Collection</i>	-
	Agregações	Objetos físicos com suas representações digitais	Objetos físicos com suas representações digitais, dados descritivos	Dados de autoridade e nomes de lugares
<b>Acesso</b>		Portal <i>SPARQL</i> API Dataset no Projeto <i>LOD-Cloud</i>	Portal DPLA API <i>Bulk download</i>	Portal <i>DigitalNZ</i> API

Fonte: Elaborado pela autora.

Evidencia-se, inicialmente, que embora o escopo dos projetos *DigitalNZ* e DPLA coincidam, esses projetos adotaram diferentes abordagens na modelagem dos dados coletados nas suas respectivas fontes: enquanto o projeto DPLA adotou um modelo de dados baseados no modelo EDM - do projeto Europeana, o *DigitalNZ* optou por se utilizar da personalização

do esquema de mapeamento de dados oferecido pelo sistema *Supplejack*. Nota-se, porém, a ausência de informações sobre as questões de modelagem e consequente descrição de dados nesse projeto, embora os responsáveis pelo sistema afirmem que estão trabalhando na construção das agregações denominadas *Concepts*.

A partir dessa proposta de agregação, destacam-se as diferentes abordagens na modelagem de dados pelos outros dois projetos, em que se observa que o projeto Europeana optou por agregar os recursos e suas representações digitais, ao passo que o projeto DPLA se aproveita dessa agregação e ainda acrescenta um modo de agrupar seus dados descritivos, por meio do uso da classe *dlpa:SourceResource*.

Tais aspectos de modelagem retomam as considerações de Glushko (2013) sobre organização, representação e recuperação de informações. Os diferentes modos de organização dos recursos culturais e de suas representações digitais - em classes e agregações - adotados por cada um dos projetos oferecem também diferentes modos de descrição e, como consequência, de recuperação de informações sobre esses recursos.

Por sua vez, as questões relativas às tecnologias utilizadas e aos meios de acesso evidenciam ainda que, embora os três projetos apresentem alternativas diversificadas de descrição e de acesso aos seus dados no âmbito do movimento *Linked Open Data*, o projeto Europeana é o único que se utiliza especificamente das ferramentas e formatos recomendados pelo consórcio W3C para os princípios *Linked Data*. O uso do formato JSON pelos outros dois projetos, porém, indica que a utilização de outras ferramentas e formatos também é válida para a estruturação *Linked Open Data* de dados e metadados, conforme demonstrado pelas diversas formas de serialização do modelo RDF no capítulo 3.

Nesse sentido, retomam-se também as discussões sobre os tipos, funções e atributos de metadados que podem ser identificados no quadro 6. De acordo com a tipologia de Riley (2017), percebe-se que os três projetos fazem uso de metadados descritivos e administrativos técnicos. O projeto DPLA, porém, se utiliza também de metadados estruturais em suas classes contextuais (*dcmitype:Collection*) e de metadados administrativos de direitos autorais (*dcterms:Rightstatement*) – assim como o projeto Europeana (*cc:Licence*).

Embora não sintetizado no quadro 6, o projeto *DigitalNZ* se destaca por utilizar metadados de uso, elencados por Pomerantz (2015). Isso porque, como supracitado, este projeto permite a recuperação dos denominados ‘sets’, formados pelas preferências de cada usuário do portal.

As funções exercidas por esses metadados refletem a sua tipologia. O capítulo 3 apresentou tais funções de acordo com Gilliland (2016), destacando que ambientes *LAM*

geralmente tem como foco a organização, descrição, busca e recuperação de recursos. O quadro 6, porém, indica que a modelagem de dados conforme os pressupostos do movimento *Linked Open Data* evidenciam outras funções exercidas por metadados no contexto desses ambientes, especialmente a criação (em ou mais formatos), contextualização e reuso de recursos.

Finalmente, a modelagem dos dados em classes e propriedades com o uso de diferentes *namespaces* pelos projetos Europeana e DPLA, bem como o esquema de metadados elaborado pelo projeto *DigitalNZ* abordam questões relativas aos atributos de metadados, elencados também por Gilliland (2016). Nesse contexto, evidenciam-se os atributos de estrutura e semântica dos metadados, traduzidos em padrões de metadados utilizados nos processos de mapeamento, descrição e acesso aos dados de coleções de patrimônio cultural, discutidos abaixo.

## **5.2 Padrões de metadados para descrição *Linked Open Data* de patrimônio cultural**

A representação de recursos envolve a descrição destes por meio do uso de metadados. O capítulo 3 apresentou discussões sobre os aspectos de estruturação de metadados em padrões de estrutura, valores, conteúdos e de formato e intercâmbio de dados, utilizados pela comunidade de patrimônio cultural para a descrição e gerenciamento de suas coleções e dos metadados gerados.

Ao mesmo tempo, as discussões sobre as características de modelagem e mapeamento de dados adotados pelas Europeana, DPLA e *DigitalNZ* evidenciaram o uso desses padrões de metadados na descrição desses objetos. Os três projetos supracitados utilizam-se do padrão de metadados estruturais *Dublin Core* e afirmam serem compatíveis com outros padrões e instrumentos como, por exemplo, o *VRA Core*, *CDWA*, *CIDOC-CRM* (DPLA, 2017a) e *Lightweight Information Describing Objects Schema* (LIDO) (EUROPEANA, 2013).

Nesse contexto, a comunidade de patrimônio cultural tem buscado adaptar seus instrumentos de representação para aplicação destes na descrição de recursos no ambiente *Web*. O quadro 7, abaixo, apresenta as soluções desenvolvidas em relação aos padrões de estrutura e de conteúdo de metadados e de formato e intercâmbio de dados. Os padrões de valores – vocabulários controlados, por sua vez, serão discutidos posteriormente.

**Quadro 7** - Desenvolvimento dos padrões de estrutura, conteúdo e de formato e intercâmbio de dados de coleções de patrimônio cultural

		<b>Padrões de formato e intercâmbio de dados</b>	<b>Soluções para descrição <i>Linked Open Data</i></b>
<b>Padrões de estrutura</b>	<i>Dublin Core Metadata Element Set</i>	<i>Simple Dublin Core XML Schema</i>	Expressão em RDF
		<i>Qualified Dublin Core XML Schema</i>	Expressão em RDF
	<i>VRA Core</i>	<i>VRA Core 4.0 XML Schema</i>	<i>VRA Core RDF Ontology</i>
<b>Padrões de conteúdo</b>	CDWA	-	CIDOC-CRM
	CCO (+ CDWA)	<i>CDWA Lite XML Schema</i>	LIDO

Fonte: Elaborado pela autora.

Como supracitado no capítulo 3, os padrões de formato e intercâmbio de dados são instrumentos de integração e interoperabilidade entre sistemas de instituições de patrimônio cultural e, por isso, podem ser considerados como soluções de interoperabilidade desenvolvidas a partir dos padrões de estrutura e de conteúdo de dados. De acordo com o apresentado no quadro 7, estes padrões ainda são os responsáveis por possibilitar o desenvolvimento de outras soluções de representação de recursos nos ambientes *Web*.

Assim, os padrões de estrutura *Dublin Core* e *VRA Core Categories* podem ser expressos no modelo RDF (DUBLIN CORE METADATA INITIATIVE, 2008; VISUAL RESOURCES ASSOCIATION, 2015). O documento ‘*Notes on DCMI specifications for expressing Dublin Core metadata in RDF*’ tem como objetivo unificar e melhorar os procedimentos de expressão em RDF de metadados no padrão *Dublin Core*, que até então contavam com diferentes diretrizes da iniciativa, dentre elas o uso do *DCMI Abstract Model* (DUBLIN CORE METADATA INITIATIVE, 2008).

A expressão em RDF do *VRA Core*, por outro lado, denomina-se *VRA Core RDF Ontology* e foi desenvolvida como uma ontologia com base no *VRA Core 4.0 XML Schema* para ser aplicado aos princípios *Linked Data* e em conjunto a outras ferramentas (VISUAL RESOURCES ASSOCIATION, 2015). De acordo com as recomendações do consórcio W3C, esta ontologia reutiliza outras ferramentas já consolidadas, em especial os *namespaces rdf*, *rdfs* e *owl* e seus elementos estão divididos em classes e propriedades - de objetos e de dados (MIXTER *et al.*, [2015]).



Em relação aos padrões de conteúdo de dados, o CDWA está sendo mapeado para integração ao CIDOC-CRM (THE GETTY RESEARCH INSTITUTE, 2017) ao mesmo tempo em que, unido ao CCO, gerou o CDWA *Lite* (THE GETTY RESEARCH INSTITUTE, 2013).

O CDWA *Lite*, por sua vez, foi incorporado ao LIDO (THE GETTY RESEARCH INSTITUTE, 2013), um XML *Schema* para *harvesting* de metadados de recursos museológicos que pode ser aplicado em conjunto a outras ferramentas, como por exemplo, o modelo de dados EDM (EUROPEANA, [201-]). Isso porque o esquema está em conformidade com normas e padrões da comunidade de patrimônio cultural, como a norma SPECTRUM e o CIDOC-CRM (ICOM INTERNATIONAL COMMITTEE FOR DOCUMENTATION, 2010).

O CIDOC-CRM trata-se de uma ontologia conceitual para dados museológicos elaborada pelo Grupo de Documentação do Comitê Internacional de Documentação (CIDOC), do Conselho Internacional de Museus (ICOM). Esta ontologia “[...] provê definições e uma estrutura formal para a descrição de conceitos e relações explícitos ou implícitos utilizados na documentação de patrimônios culturais” (ICOM/CIDOC, [201-], *online*, tradução nossa<sup>33</sup>).

Sendo uma ferramenta conceitual, pode ser utilizada de diferentes modos, como por exemplo, em conjunto ao próprio modelo RDF e à linguagem XML (ICOM/CIDOC, 2013). Para Santos (2016), constitui-se em uma referência multidisciplinar, pois tem como finalidades a modelagem de aplicações, a interoperabilidade/integração de dados e o processamento de consultas complexas que necessitam de inferências. Nesse sentido, pode ser utilizada por instituições de patrimônio cultural que queiram iniciar o mapeamento de integração de dados de suas coleções.

Nesse contexto, ressaltam-se os esforços relacionados às ontologias, aos vocabulários de dados e aos vocabulários controlados utilizados pela comunidade de patrimônio cultural. A iniciativa *Linked Open Vocabularies* (LOV) (2012, *online*, tradução nossa<sup>34</sup>) explica na sua documentação que

---

<sup>33</sup> *The CIDOC Conceptual Reference Model (CRM) provides definitions and a formal structure for describing the implicit and explicit concepts and relationships used in cultural heritage documentation*

<sup>34</sup> *We must first clarify the meaning of "Vocabulary" in the context of this document. We do not stick to the definition of vocabulary in linguistics meaning but approach it in terms of its use in the Semantic Web. In this context, a vocabulary is basically synonymous of ontology. However, we differentiate vocabulary from an ontology by characteristics enabling reuse and integration by other vocabularies:*

- *Small size;*
- *Low formal constraints (basically RDFS and a fistful of OWL);*

Devemos primeiro esclarecer o significado de ‘vocabulário’ [...]. Nós não nos atemos à definição de vocabulário em significado linguístico, mas o abordamos em termos de seu uso na *Web Semântica*. Neste contexto, um vocabulário é basicamente sinônimo de ontologia. No entanto, diferenciamos o vocabulário de uma ontologia por características que permitem a reutilização e a integração por outros vocabulários: tamanho reduzido, poucas restrições formais (basicamente RDF-S e um pouco de OWL), poucas instâncias (exceto para exemplos) e documentação rica para o usuário (etiquetas, comentários, definição, descrição, etc.).

No contexto geral da *Web*, então, esta iniciativa tem como objetivo fomentar a utilização de vocabulários de dados e a construção ou extensão de novos, de acordo com as necessidades de publicadores e de usuários de dados (LOV, [201-]). Dentre os vocabulários de domínio da comunidade de patrimônio cultural disponibilizados por esta iniciativa estão a própria ontologia CIDOC-CRM e o modelo de dados EDM, bem como a expressão em RDF dos conceitos e relacionamentos dos *Functional Requirements for Bibliographic Records* (FRBR) e o vocabulário do BIBFRAME.

Os FRBR constituem um modelo conceitual elaborado pela *International Federation of Library Associations and Institutions* (IFLA) para estruturação de registros bibliográficos de acordo com modelos entidade-relacionamento (FUSCO, 2010), enquanto o BIBFRAME constitui-se em um padrão de metadados elaborado pela *Library of Congress* que se utiliza dos princípios *Linked Data* e também dos FRBR para a descrição e conexão de recursos do domínio bibliográfico (ARAKAKI, 2016).

Especificamente no domínio bibliográfico, destacam-se ainda os esforços do OCLC, líder do consórcio VIAF. Este consórcio tem como um de seus objetivos a publicação de arquivos de dados de autoridades em *Linked Open Data* (ROMANETTO; SANTOS; ALVES, 2017). Nesse sentido, os dados do VIAF, assim como a ontologia *Schema.org*, são utilizados na modelagem e posterior publicação *Linked Open Data* dos dados do catálogo *WorldCat*, constituindo o denominado *WorldCat Linked Data Vocabulary* (OCLC, 2015).

Finalmente, o quadro 8 apresenta as diferentes soluções desenvolvidas para a utilização dos vocabulários controlados, elencados no capítulo 3, na descrição *Linked Open Data* de coleções de patrimônio cultural.

- 
- *Few instances except for examples;*
  - *Rich user documentation (Labels, comments, definition, description, etc.)*

**Quadro 8** - Soluções para uso dos padrões de valores de dados na descrição *Linked Open Data* de coleções de patrimônio cultural

Padrões de valores de dados		Soluções para descrição <i>Linked Open Data</i>
TGM		URI para cada um dos seus campos e valores de dados <i>Dataset</i> publicado (formato XML, JSON e outros)
Vocabulários <i>Getty</i>	AAT	URI para cada um dos seus termos, hierarquia e mapeamento em RDF Mapeamento em RDF
	TGN	URI para cada um dos seus termos, hierarquia e mapeamento em RDF Mapeamento em RDF
	ULAN	URI para cada um dos seus termos, hierarquia e mapeamento em RDF Mapeamento em RDF
	CONA	URI para cada um dos seus termos e hierarquia CIDOC-CRM
ICONCLASS		URI para cada notação Expressão em RDF e JSON

Fonte: Elaborado pela autora.

A *Library of Congress* apresenta em seu *site* uma série de serviços voltados à expressão de seus padrões e vocabulários de acordo com princípios *Linked Data*. Nestes serviços, cada um dos campos do vocabulário controlado e cada um dos seus valores de dados recebem um URI resolvível, ao mesmo tempo em que cada vocabulário está publicado como um *dataset* e disponível para *download* nos formatos XML, JSON e outros (LIBRARY OF CONGRESS, [20--]).

Dentre estes vocabulários está o TGM, podendo-se destacar também os vocabulários relativos à preservação digital de recursos, que incluem termos sobre o *status* do *copyright* e sobre a categoria do objeto descrito. Instituições de patrimônio cultural podem, então, recorrer a esses serviços para descrever diferentes tipos de objetos que não apenas recursos bibliográficos.

No domínio museológico, os vocabulários *Getty* também estão disponíveis para serem utilizados com os princípios *Linked Data*, ao mesmo tempo em que o instituto tem trabalhado na publicação desses vocabulários no escopo do movimento *Linked Open Data* (THE GETTY RESEARCH INSTITUTE, 2017). Assim, os vocabulários AAT, TGN, ULAN possuem URIs - legíveis por humanos e máquinas - para os registros completos de cada um dos seus termos,

para sua hierarquia e para o mapeamento em RDF (THE GETTY RESEARCH INSTITUTE, 2017).

O vocabulário CONA, porém, é uma exceção em relação à modelagem RDF, uma vez que também está sendo mapeado pelo CIDOC-CRM (THE GETTY RESEARCH INSTITUTE, 2017). Desse modo, possui URIs apenas para seus registros de termos e para sua estrutura hierárquica.

Assim como os vocabulários *Getty*, o sistema de classificação ICONCLASS possui um URI para cada notação, que pode ser resolvido em uma página HTML ou ser expresso em RDF ou JSON, e ainda permite a recuperação dos textos explicativos de cada notação (ICONCLASS, 2015). Percebe-se, então, a possibilidade de instituições de patrimônio cultural utilizarem as informações contidas nos sistemas de classificação para enriquecimento de descrições de recursos em ambientes *Linked Open Data*.

Considerando-se todas essas contribuições, nota-se que os padrões de estrutura, valores, conteúdos e de formato e intercâmbio de metadados de patrimônio cultural apresentados são aplicáveis ao movimento *Linked Open Data* em diferentes contextos.

Destacam-se primeiramente os padrões de intercâmbio e formato de dados, uma vez que o uso da linguagem XML e do XML *Schema* permite que cada um dos outros tipos de padrões sejam expressos no modelo RDF ou aplicados em outros procedimentos que buscam a interoperabilidade entre sistemas. Em contraposição, percebe-se que tais padrões não possuem aderência ao formato JSON, o que limita suas possíveis aplicações em contextos e ambientes que se utilizam desse formato.

Os padrões de valores, notadamente os vocabulários, também se destacam, pois são objeto de esforços de organizações como *Library of Congress* e *The Getty Research Institute* para sua integração com ontologias, vocabulários de dados e modelos de dados, ao mesmo tempo em que permitem a sua livre utilização no enriquecimento de descrições de coleções de patrimônio cultural, por meio da disponibilização de URIs para cada um de seus termos.

No escopo das Humanidades Digitais, percebem-se as diversas possibilidades de aplicação das tecnologias e ferramentas para a promoção do acesso às informações culturais pelas instituições de patrimônio cultural. As iniciativas apresentadas no capítulo 4, em especial os projetos Europeana, DPLA e *DigitalNZ*, apontaram algumas dessas possibilidades, compatíveis com o escopo de atuação e com as realidades institucionais de cada uma delas.

Nesse contexto, a utilização de metadados padronizados e a representação digital de recursos configuram ações de preservação do patrimônio cultural que, quando alinhados aos pressupostos do movimento *Linked Open Data*, oferecem modos de organização,

representação e acesso sustentáveis a esse patrimônio. Isso porque, como supracitado, a estruturação *Linked Open Data* atua diretamente em questões de interoperabilidade técnica e semântica entre os sistemas de instituições culturais. Assim, tais pressupostos ampliam a possibilidade de descoberta de recursos para além dos metadados descritivos, ao mesmo tempo em que permitem o efetivo uso e reuso de dados estruturados.

As instituições *LAM* podem, então, optar por diferentes caminhos nos processos de descrição *Linked Open Data* de suas coleções, opções essas que se ajustem aos seus ambientes institucionais, incluindo as ferramentas e instrumentos já consolidados nos seus processos de representação.

## 6 CONSIDERAÇÕES FINAIS

A explosão informacional que deu origem à Ciência da Informação e o desenvolvimento do ambiente digital - em especial da plataforma *Web* para uma *Web Semântica* - trazem desafios e, ao mesmo tempo, novas possibilidades de se pensar a organização, a representação, acesso, uso e reuso dos recursos informacionais. No âmbito das Humanidades Digitais, consideram-se as possibilidades de promoção do acesso à informação cultural por meio do uso das ferramentas oferecidas pelo movimento *Linked Open Data* para a representação do patrimônio cultural.

Desse modo, essa pesquisa teve como objetivo geral analisar o *Linked Open Data* em instituições de patrimônio cultural, por meio de levantamento bibliográfico e documental acerca da questão de pesquisa. Este objetivo geral desdobrou-se em objetivos específicos. O primeiro deles buscou identificar os relacionamentos do patrimônio cultural com a Ciência da Informação e as Humanidades Digitais. Desse modo, foram discutidas questões relativas ao patrimônio cultural, suas relações com a Ciência da Informação e com as Humanidades Digitais. Percebeu-se que estes dois campos tratam do acesso à informação cultural registrada nas coleções de patrimônio cultural, que este possui características materiais e imateriais, e que pode fazer parte de um novo tipo de patrimônio, o patrimônio digital. Verificou-se também que o patrimônio cultural não está presente apenas nos museus, mas também em bibliotecas e em arquivos.

Nesse sentido, apresentaram-se considerações a respeito do campo da organização e representação da informação nos ambientes digitais, a partir da exposição dos conceitos de dados e metadados e da identificação das características, tipos e funções de metadados, das quais se destacou o aspecto da estruturação de metadados por meio de padrões. Foram elencados então, os padrões de metadados utilizados pela comunidade de patrimônio cultural para a representação de suas coleções de acordo com a tipologia de Gilliland (2016): padrões de estrutura, valores e conteúdo de metadados e de formato e intercâmbio de dados.

As discussões foram, então, direcionadas ao segundo objetivo específico: conceituar os princípios *Linked Data* e *Open Data*, o movimento *Linked Open Data* e as ferramentas de descrição de recursos no ambiente *Web*. Nesse contexto, percebeu-se a dinâmica entre princípios, movimentos e iniciativas: princípios tratam de requisitos básicos e concretos a serem observados, cuja aplicação prática é incentivada pelos movimentos; estes, por sua vez, se concretizam em iniciativas e projetos liderados por instituições ou de grupos de pessoas.

Assim, o movimento *Linked Open Data* se utiliza dos princípios *Open Data* e *Linked Data* para promover acesso aos dados estruturados provenientes de diversos domínios, dentre eles, dados de coleções de patrimônio cultural. Foram apresentadas, então, as ferramentas URI, RDF e XML, que tratam da descrição de recursos no âmbito da plataforma *Web* e também do movimento *Linked Open Data*.

Tendo em vista a possibilidade de inserção de instituições de patrimônio cultural no escopo desse movimento, buscou-se pontuar iniciativas *Linked Open Data* que possuem essa finalidade, de acordo com o terceiro objetivo específico proposto para a pesquisa. Foram, então, apresentadas iniciativas *Linked Open Data* em instituições de patrimônio cultural em âmbito internacional. As contribuições institucionais em âmbito nacional, porém, indicam que as instituições culturais brasileiras têm trabalhado principalmente com os princípios *Open Data* para a disponibilização de dados administrativos e governamentais.

As contribuições acadêmicas a respeito do uso das ferramentas do movimento *Linked Open Data* para a descrição e disponibilização de dados de coleções de patrimônio cultural evidenciaram preocupações relativas à sustentabilidade dos processos de descrição e de acesso aos dados dessas coleções, englobando todo o ciclo de gerenciamento informacional. Assim, mostram-se emergentes as práticas para a manutenção dos dados orientados às questões sustentáveis, com o foco nas ações da instituição como um todo e de forma sistêmica: não há como disseminar e prover acesso a um conteúdo que não está devidamente representado, da mesma forma que não há razão de se representar algo que não será mostrado e disseminado.

Conforme a proposta do quarto objetivo específico da pesquisa, os projetos Europeana, DPLA e *DigitalNZ* foram selecionados e explorados, pois fornecem acesso aos seus dados estruturados por meio de APIs. Assim, os modelos dados adotados pelos projetos Europeana e DPLA e o esquema de dados utilizado pelo projeto *DigitalNZ* foram apresentados e discutidos sob o ponto de vista do escopo de atuação dessas iniciativas, das tecnologias e ferramentas utilizadas e da própria configuração dos seus modelos e esquemas de dados. Cada um desses projetos se destacou de modo particular, pois optaram por procedimentos que condizem com o seu escopo de atuação e a sua realidade institucional.

Finalmente, de acordo com o último objetivo específico, apresentaram-se as soluções desenvolvidas pela comunidade de patrimônio cultural para uso dos seus padrões de metadados na descrição *Linked Open Data* de suas coleções. Dentre os tipos de padrões utilizados por essa comunidade, destacaram-se os padrões de formato e intercâmbio de dados, que permitem a descrição *Linked Open Data* por meio do uso da linguagem XML e do

modelo RDF, bem como os padrões de valores de dados, que se utilizam de URIs para identificação de seus termos e podem ser utilizados livremente para enriquecer a descrição de coleções de patrimônio cultural. Embora essas três ferramentas façam parte dos princípios *Linked Data*, percebeu-se também o uso do formato JSON pelas iniciativas DPLA e *DigitalNZ*, bem como em algumas das soluções relativas aos padrões de metadados apresentadas. Este formato configura como outra opção de serialização do modelo RDF que pode ser aplicada para a descrição *Linked Open Data* de coleções de patrimônio cultural.

Diante desses resultados, considera-se que a comunidade de patrimônio cultural tem buscado se aproveitar dos benefícios oferecidos pelo uso das ferramentas do movimento *Linked Open Data* por meio da adaptação de seus instrumentos de representação para a realidade do ambiente *Web*. Ao mesmo tempo, nota-se que as instituições de patrimônio cultural buscam se aproximar das práticas e objetivos propostos pelas Humanidades Digitais por meio de suas iniciativas *Linked Open Data* de disponibilização de coleções culturais.

Os mesmos resultados ressaltam a relevância que o movimento *Linked Open Data* possui em relação aos aspectos de interoperabilidade técnica e semântica dos sistemas de recuperação de informações, tema fundamental para a Ciência da Informação e, de modo mais amplo, para as Humanidades Digitais.

Ressaltam-se limitações de execução da pesquisa. Esta última envolveu aspectos da literatura documental sobre os projetos *Linked Open Data* apresentados: a iniciativa *Google Arts & Culture* não disponibiliza documentação sobre o modo como seus dados estão descritos ou sobre como a iniciativa oferece seus serviços, ao mesmo tempo em que o projeto *DigitalNZ*, como supracitado, possui documentação com linguagem técnica que dificulta seu entendimento por usuários e profissionais leigos e conseqüentemente, compromete a customização oferecida pelo seu sistema à esses mesmos usuários. Em âmbito nacional, o projeto Tainacan também possui pouca documentação a ser explorada.

Em conjunto às limitações de escopo dessa pesquisa, estes fatores oferecem perspectivas para outras pesquisas. Dentre essas possibilidades encontram-se estudos que se voltem aos aspectos específicos da publicação de dados conforme os princípios *Open Data* e as diretrizes do consórcio W3C, expressas no supracitado documento '*Data on the Web best practices*'. Tais estudos poderiam aprofundar as questões de publicação de dados de autoridade ou discutir a publicação de outros tipos de dados como, por exemplo, os dados de vocabulários controlados. Poderiam ainda tratar especificamente dos aspectos de sustentabilidade na publicação de dados ou suas relações com a curadoria digital.



Outras possibilidades dizem respeito ao aprofundamento dos estudos sobre os vocabulários de dados e as ontologias desenvolvidas para a representação de coleções de patrimônio cultural na *Web*, dentre eles os citados nessa pesquisa: o vocabulário do BIBFRAME, os FRBR e a ontologia CIDOC-CRM. Em se tratando especificamente do patrimônio cultural, apontam-se também as possibilidades de estudos que tratam do patrimônio cultural digital e seus aspectos intangíveis no ambiente *Web*.

No aspecto prático, destacam-se ainda possibilidades relativas à proposição de métodos e procedimentos de descrição e publicação de dados de coleções de patrimônio cultural ou de outros tipos de dados pertinentes aos ambientes *LAM*. No mesmo sentido, o uso do formato JSON por algumas das iniciativas e soluções apresentadas indica que os esforços da comunidade de patrimônio cultural podem ser complementados com a inclusão de outras ferramentas de descrição *Linked Open Data* no rol de opções para essa comunidade.

Todas essas propostas buscam o aprofundamento da comunidade acadêmica e profissional na temática da representação e acesso ao patrimônio cultural e podem proporcionar caminhos que auxiliem as instituições de patrimônio cultural a se inserirem no movimento *Linked Open Data* e, de modo geral, nas Humanidades Digitais.

## REFERÊNCIAS

- ALMEIDA, M. A.; DAMIAN, I. P. M. Humanidades Digitais: um campo praxiológico para mediações e políticas culturais?. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 16, 2015, João Pessoa. **Anais...** João Pessoa, Universidade Federal da Paraíba, 2015. Disponível em: <<http://repositorios.questoesemrede.uff.br/repositorios/handle/123456789/2838>>. Acesso em: 20 jun. 2018.
- ALMEIDA, M. B. Uma introdução ao XML, sua utilização na Internet e alguns conceitos complementares. **Ciência da Informação**, Brasília, v. 31, n. 2, p. 5–13, ago. 2002. Disponível em: <[http://www.scielo.br/scielo.php?pid=s010019652002000200001&script=sci\\_abstract&lng=pt](http://www.scielo.br/scielo.php?pid=s010019652002000200001&script=sci_abstract&lng=pt)>. Acesso em 24 jan. 2017.
- ALVES, R. C. V. **Metadados como elementos do processo de catalogação**. 2010. 132 f. Tese (Doutorado em Ciência da Informação) – Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2010. Disponível em: <<http://repositorio.unesp.br/handle/11449/103361>>. Acesso em: 22 ago. 2017.
- ALVES, R. C. V. **Web semântica: uma análise focada no uso de metadados**. 2005. 180 f. Dissertação (Mestrado em Ciência da Informação) - Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2005.
- ARAKAKI, F. A. **Linked Data: ligação de dados bibliográficos**. 2016. 146 f. Dissertação (Mestrado em Ciência da Informação) - Universidade Estadual Paulista, Marília, SP, 2016. Disponível em: <[http://repositorio.unesp.br/bitstream/handle/11449/147979/arakaki\\_fa\\_me\\_mar.pdf?sequence=2&isAllowed=y](http://repositorio.unesp.br/bitstream/handle/11449/147979/arakaki_fa_me_mar.pdf?sequence=2&isAllowed=y)>. Acesso em: 12 jan. 2017.
- ARARIPE, F. M. A. Do patrimônio cultural e seus significados. **Transinformação**, Campinas, v. 16, n. 2, p. 111–122, ago. 2004. Disponível em <[http://www.scielo.br/scielo.php?script=sci\\_arttext&pid=S010337862004000200001&lng=en&nrm=iso](http://www.scielo.br/scielo.php?script=sci_arttext&pid=S010337862004000200001&lng=en&nrm=iso)>. Acesso em: 13 abr. 2017.
- ARAÚJO, C. A. Á. Fundamentos da Ciência da Informação: correntes teóricas e o conceito de informação. **Perspectivas em Gestão & Conhecimento**, João Pessoa, v. 4, n. 1, p. 57–79, 2014.
- ASSUMPÇÃO, F. S. **Conversão de registros em XML para MARC 21: um modelo baseado em folhas de estilo XSLT**. 2013. 135 f. Dissertação (mestrado) - Universidade Estadual Paulista Júlio de Mesquita Filho, Faculdade de Filosofia e Ciências de Marília, 2013. Disponível em: <<http://repositorio.unesp.br/handle/11449/93658>>. Acesso em: 06 set. 2017.

ASSUMPÇÃO, F. S. **Modelo para a publicação de dados de autoridade como Linked Data**. 2018. 208 f. Tese (Doutorado em Ciência da Informação) - Universidade Estadual Paulista Júlio de Mesquita Filho, Faculdade de Filosofia e Ciências de Marília, 2018.

BERNERS-LEE, T. **Linked Data: design Issues**. 2006. Disponível em: <<https://www.w3.org/DesignIssues/LinkedData.html>>. Acesso em: 4 dez. 2016.

BERNERS-LEE, T. **Semantic Web road map**. 1998. Disponível em: <<http://www.w3.org/DesignIssues/Semantic.html>>. Acesso em: 17 mai 2017.

BIZER, C.; HEATH, T.; BERNERS-LEE, T. Linked data: the story so far. **International Journal on Semantic Web and Information Systems**, EUA, v. 5, n. 3, p. 1-22, 2009. Disponível em: <<http://tomheath.com/papers/bizer-heath-berners-lee-ijswis-linked-data.pdf>>. Acesso em: 13 abr. 2017.

BRASIL. Constituição da República Federativa do Brasil de 1988. 1988. Disponível em: <[http://www.planalto.gov.br/ccivil\\_03/constituicao/constituicao.htm](http://www.planalto.gov.br/ccivil_03/constituicao/constituicao.htm)>. Acesso em: 17 de mar. 2017.

BRASIL. Decreto-lei nº 25, de 30 de novembro de 1937. Organiza a proteção do patrimônio histórico e artístico nacional. **Diário Oficial da União**, Rio de Janeiro, 06 dez. 1937, Seção 1. Disponível em: <[http://portal.iphan.gov.br/uploads/legislacao/Decreto\\_no\\_25\\_de\\_30\\_de\\_novembro\\_de\\_1937.pdf](http://portal.iphan.gov.br/uploads/legislacao/Decreto_no_25_de_30_de_novembro_de_1937.pdf)>. Acesso em: 04 jul. 2017.

BRASIL. Decreto nº 3551, de 04 de agosto de 2000. Institui o Registro de Bens Culturais de Natureza Imaterial que constituem patrimônio cultural brasileiro, cria o Programa Nacional do Patrimônio Imaterial e dá outras providências. **Diário Oficial da República Federativa do Brasil**, Poder Executivo, Brasília, DF, 07 ago. 2000, Seção 1.

BRASIL. Ministério da Cultura (MinC). Plano de dados abertos do MinC e órgãos vinculados. 2017. Disponível em: <<http://plano.dados.cultura.gov.br/>>. Acesso em: 03 out. 2017.

BRASIL. Ministério da Cultura (MinC). Transparência: MinC lança portal Plano de Dados Abertos, 31 de maio de 2017. Disponível em: <[http://www.cultura.gov.br/noticias-destaques/-/asset\\_publisher/OiKX3xlR9iTn/content/transparencia-minc-lanca-portal-plano-de-dados-abertos/10883](http://www.cultura.gov.br/noticias-destaques/-/asset_publisher/OiKX3xlR9iTn/content/transparencia-minc-lanca-portal-plano-de-dados-abertos/10883)>. Acesso em: 03 out. 2017.

BRASIL. Ministério da Justiça e Segurança Pública. Arquivo Nacional. Plano de dados abertos – Biênio 2017-2019. In: \_\_\_\_\_. Portaria nº 284, de 30 de junho de 2017. **Boletim interno especial**. Disponível em: <[http://www.arquivonacional.gov.br/images/Anexos\\_definitivos/Boletim\\_Interno\\_Especial\\_n%C2%BA\\_13-2017.pdf](http://www.arquivonacional.gov.br/images/Anexos_definitivos/Boletim_Interno_Especial_n%C2%BA_13-2017.pdf)>. Acesso em: 03 out. 2017.

BREITMAN, K. *et al.* **Publicação de Dados Governamentais no Padrão Linked Data.** PUC Rio de Janeiro, 2003. Disponível em: <<http://www.w3c.br/cursos/dados-abertos/curso/Parte-1-Modulo-3-Web%20Semantica.pdf>>. Acesso em: 7 jun. 2018.

BUCKLAND, M. K. Information as thing. **Journal of the American Society for Information Science (1986-1998)**, EUA, v. 42, n. 5, p. 351, 1991.

BYRNE, G.; GODDARD, L. The Strongest Link: Libraries and Linked Data. **D-Lib Magazine, online**, v. 16, n. 11/12, 2010.

CAGNAZZO, L. **Linked Data: implementation, use, and perceptions across European National Libraries.** 2017. 129 f. Dissertação (Mestrado em Ciência da Informação)—Glasgow, UK: University of Strathclyde, 2017.

CAPURRO, Rafael. **Epistemologia e Ciência da Informação.** Tradução: Ana Maria Rezende Cabral, Eduardo Wense Dias, Isis Paim, Ligia Maria Moreira Dumont, Marta Pinheiro Aun e Mônica Erichsen Nassif Borges. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 5, 2003, Belo Horizonte. Disponível em: <[http://www.capurro.de/enancib\\_p.htm](http://www.capurro.de/enancib_p.htm)>. Acesso em: 09 maio 2017.

CAPURRO, R.; HJORLAND, B. O conceito de Informação. Tradução: Ana Maria Pereira Cardoso, Maria da Glória Achtschin Ferreira, Marco Antônio de Azevedo. **Perspectivas em Ciência da Informação**, v. 12, n.1, p.148-207, jan./abr. 2007.

CASTELLS, M. **A sociedade em rede.** 10 ed. São Paulo: Paz e Terra, 2007. 698 p. (A Era da Informação: economia, sociedade e cultura, v.1).

CATARINO, M. E.; SOUZA, T. B. A representação descritiva no contexto da web semântica. **Transinformação**, Campinas, v. 24, n. 2, p. 77-90, ago. 2012. Disponível em: <[http://www.scielo.br/scielo.php?script=sci\\_arttext&pid=S010337862012000200001&lng=en&nrm=iso](http://www.scielo.br/scielo.php?script=sci_arttext&pid=S010337862012000200001&lng=en&nrm=iso)>. Acesso em: 13 abr. 2017.

CAVALCANTE, L. E. Patrimônio digital e informação: política, cultura e diversidade. **Encontros Bibli: revista eletrônica de biblioteconomia e ciência da informação**, Florianópolis, v. 12, n. 23, p. 152-170, dez. 2007. Disponível em: <<https://periodicos.ufsc.br/index.php/eb/article/view/1518-2924.2007v12n23p152/401>>. Acesso em: 05 jul. 2017.

COLLECTIONS TRUST. **Introduction to Spectrum 5.0.** 2017. Disponível em: <<https://collectionstrust.org.uk/spectrum/spectrum-5/>>. Acesso em: 23 maio. 2018.

COMITÉ DO PATRIMÓNIO MUNDIAL. **Declaração de Budapeste sobre o Património Mundial.** Comité do Património Mundial, 2002. Disponível em:

<<http://www.patrimoniocultural.gov.pt/media/uploads/cc/declaracaoBudapestesobrepatrimoniomundial2002.pdf>>. Acesso em: 8 jun. 2018

CONEGLIAN, C. S.; SANTAREM SEGUNDO, J. E. Europeana no Linked Open Data: conceitos de Web Semântica na dimensão aplicada das Humanidades Digitais. **Encontros Bibli: revista eletrônica de biblioteconomia e ciência da informação**, Florianópolis, v. 22, n. 48, p. 88-99, jan./abr. 2017. Disponível em: <<https://periodicos.ufsc.br/index.php/eb/article/view/1518-2924.2017v22n48p88>>. Acesso em: 20 jun. 2018.

CONGRESSO INTERNACIONAL EM HUMANIDADES DIGITAIS (ED.). **HDRio2018**. Anais do I Congresso Internacional em Humanidades Digitais. **Anais...** In: I CONGRESSO INTERNACIONAL EM HUMANIDADES DIGITAIS. Rio de Janeiro: FGV Eventos, 2018. Disponível em: <<http://eventos.fgv.br/hdrio2018>>. Acesso em: 7 jun. 2018

COSTA, A. A. P. Uma proposta de Linked Data de acervos em uma rede de museus. **Projetos e Dissertações em Sistemas de Informação e Gestão do Conhecimento**, v. 5, n. 2, 2017.

COUNCIL ON LIBRARY AND INFORMATION RESOURCES. **Building and sustaining digital collections: models for libraries and museums**. Washington, DC: Council on Library and Information Resources, 2001. Disponível em: <[www.clir.org/pubs/reports/pub100/pub100.pdf](http://www.clir.org/pubs/reports/pub100/pub100.pdf)>. Acesso em: 13 mar. 2017.

CREATIVE COMMONS. **O que é Creative Commons?** [20--]. Disponível em: <<https://br.creativecommons.org/sobre/>>. Acesso em: 28 fev. 2018.

DACOS, M. **Manifeste des Digital humanities –THATCamp Paris**, 2011. Disponível em: <<https://tcp.hypotheses.org/category/manifeste>>. Acesso em: 8 jun. 2018

DAHLSTRÖM, M.; HANSSON, J.; KJELLMAN, U. As we may digitize. **LIBER Quarterly**, v. 21, n. 3–4, p. 455, 11 abr. 2012, p. 455-474. Disponível em: <<https://www.liberquarterly.eu/articles/10.18352/lq.8036/>>. Acesso em 14 mar. 2017.

DI GIORGIO, S. Culturaitalia, the Italian National Content Aggregator in Europeana. **Procedia Computer Science**, *online*, v. 38, p. 40–43, 2014. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S1877050914013684>>. Acesso em: 29 set. 2017.

DIGITAL NEW ZEALAND (DIGITALNZ). **About DigitalNZ**. 2017. Disponível em: <<https://digitalnz.org/about>>. Acesso em: 8 jun. 2018.

DIGITAL NEW ZEALAND (DIGITALNZ). **API docs v3**. [201-]. Disponível em: <<https://digitalnz.org/developers/api-docs-v3>>. Acesso em: 30 mai. 2018.

DIGITAL NEW ZEALAND (DIGITALNZ). **Current content partners**. [201-]. Disponível em: <<https://digitalnz.org/partners/current-content-partners>>. Acesso em: 31 mai. 2018.

DIGITAL NEW ZEALAND (DIGITALNZ). **DigitalNZ Metadata Dictionary**. [201-]. Disponível em: <[https://docs.google.com/document/pub?id=1Z3I\\_ckQWjnQQ4SzpORbClcIXUheO-Jd4jt-oZFuMcoQ](https://docs.google.com/document/pub?id=1Z3I_ckQWjnQQ4SzpORbClcIXUheO-Jd4jt-oZFuMcoQ)>. Acesso em: 21 fev. 2018.

DIGITAL NEW ZEALAND (DIGITALNZ). **Our history**. [2017]. Disponível em: <<https://digitalnz.org/about/our-history>>. Acesso em: 20 fev. 2018.

DIGITAL NEW ZEALAND (DIGITALNZ). **Supplejack**. [2017]. Disponível em: <<https://digitalnz.org/developers/supplejack>>. Acesso em: 20 fev. 2018.

DIGITAL PUBLIC LIBRARY OF AMERICA (DPLA). **API Basics**. [201-]. Disponível em: <<https://pro.dp.la/developers/api-basics/>>. Acesso em: 01 jun. 2018.

DIGITAL PUBLIC LIBRARY OF AMERICA (DPLA). **Bulk download**. [201-]. Disponível em: <<https://pro.dp.la/developers/bulk-download>>. Acesso em: 09 jun. 2018.

DIGITAL PUBLIC LIBRARY OF AMERICA (DPLA). **Introduction to the DPLA Metadata Application Profile, version 5.0**. 2017. Disponível em: <<https://drive.google.com/file/d/1kMxXgFrGwu3i7LBLFOj6VZRQFuQzqkHk/view>>. Acesso em: 03 abr. 2018.

DIGITAL PUBLIC LIBRARY OF AMERICA (DPLA). **Metadata application profile, version 5.0**. 2017. Disponível em: <[https://drive.google.com/file/d/1fJEWhnYy5Ch7\\_ef\\_V48-FAViA72OieG/view](https://drive.google.com/file/d/1fJEWhnYy5Ch7_ef_V48-FAViA72OieG/view)>. Acesso em: 09 jun. 2018.

DODEBEI, V. Patrimônio e memória digital. **Revista Morpheus-Estudos Interdisciplinares em Memória Social**, *online*, v. 5, n. 8, 2006.

DOUEIHI, M. **La gran conversión digital**. México: Fondo de cultura económica, 2010.

DUBLIN CORE METADATA INICIATIVE. 2018. Disponível em: <<http://dublincore.org>>. Acesso em: 21 jun. 2018.

DUBLIN CORE METADATA INICIATIVE. **DCMI Metadata Terms**. 2017. Disponível em: <<http://dublincore.org/documents/dcmi-terms/>>. Acesso em: 8 jun. 2018.

DUBLIN CORE METADATA INICIATIVE (DCMI). **Notes on DCMI specifications for expressing Dublin Core metadata in RDF**. 2008. Disponível em: <<http://dublincore.org/documents/dc-rdf-notes/>>. Acesso em: 03 abr. 2018.

DUBLIN CORE METADATA INITIATIVE (DCMI). **XML schemas to support the Guidelines for implementing Dublin Core in XML**. 2017. Disponível em: <<http://dublincore.org/schemas/xmls/>>. Acesso em: 04 set. 2017.

EUROPEANA. **Europeana Data Model Documentation**. 2014. Disponível em: <<https://pro.europeana.eu/page/edm-documentation>>. Acesso em: 28 fev. 2018

EUROPEANA. **Europeana Data Model Primer**, 2013. Disponível em: <[https://pro.europeana.eu/files/Europeana\\_Professional/Share\\_your\\_data/Technical\\_requirements/EDM\\_Documentation/EDM\\_Primer\\_130714.pdf](https://pro.europeana.eu/files/Europeana_Professional/Share_your_data/Technical_requirements/EDM_Documentation/EDM_Primer_130714.pdf)>. Acesso em: 8 jun. 2018.

EUROPEANA. **Europeana Data Model: mapping guidelines v.2.4**. 2017. Disponível em: <[https://pro.europeana.eu/files/Europeana\\_Professional/Share\\_your\\_data/Technical\\_requirements/EDM\\_Documentation/EDM\\_Mapping\\_Guidelines\\_v2.4\\_102017.pdf](https://pro.europeana.eu/files/Europeana_Professional/Share_your_data/Technical_requirements/EDM_Documentation/EDM_Mapping_Guidelines_v2.4_102017.pdf)>. Acesso em: 23 fev. 2018.

EUROPEANA. **History**. 2017. Disponível em: <<https://pro.europeana.eu/our-mission/history>>. Acesso em: 02 out. 2017.

EUROPEANA. **Introduction**. 2018. Disponível em: <<https://pro.europeana.eu/page/linked-open-data>>. Acesso em: 21 jun. 2018.

EUROPEANA. **The Europeana Data Model for cultural heritage**. [201-]. Disponível em: <[https://pro.europeana.eu/files/Europeana\\_Professional/Share\\_your\\_data/Technical\\_requirements/EDM\\_Documentation/EDM\\_Factsheet.pdf](https://pro.europeana.eu/files/Europeana_Professional/Share_your_data/Technical_requirements/EDM_Documentation/EDM_Factsheet.pdf)>. Acesso em: 23 fev. 2018.

FEOFILOFF, P.; KOHAYAKAWA, Y.; WAKABAYASHI, Y. **Uma introdução sucinta à teoria dos grafos**. Disponível em: <<https://www.ime.usp.br/~pf/teoriadosgrafos/>>. Acesso em: 8 jun. 2018.

FERREIRA, J. A. **Wikis semânticos: da Web para a Web Semântica**. 2014. 131 f. Dissertação (Mestrado em Ciência da Informação). Faculdade de Filosofia e Ciências – Universidade Estadual Paulista, Marília, 2014. Disponível em: <<http://repositorio.unesp.br/handle/11449/108380>>. Acesso em: 06 set. 2017.

FERREIRA, J. A.; SANTOS, P. L. V. A. DA C. O modelo de dados Resource Description Framework (RDF) e o seu papel na descrição de recursos. **Informação & Sociedade**, João Pessoa, v. 23, n. 2, 2013.

FREIRE, G. Ciência da informação: temática, histórias e fundamentos. **Perspectivas em Ciência da Informação**, Belo Horizonte, v.11 n.1, p. 6-19, jan./abr. 2006.

FONSECA, M. C. L. The registry of intangible heritage: the Brazilian experience. **Museum International: Intangible Heritage**, Paris, Unesco, v. 56, n. 1-2, p. 166- 172, 2004.



Disponível em: <<http://unesdoc.unesco.org/images/0013/001358/135852e.pdf#135853>>.  
Acesso em: 21 jun. 2018.

FUSCO, E. **Modelos conceituais de dados como parte do processo da catalogação:** perspectiva de uso dos FRBR no desenvolvimento de catálogos bibliográficos digitais. 2010. 249 f. Tese (Doutorado em Ciência da Informação)–Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2010. Disponível em: <<http://repositorio.unesp.br/handle/11449/103369>>. Acesso em: 25 mai. 2018.

GERHARDT, T. E.; SILVEIRA, D. T. (Orgs). **Métodos de pesquisa**. Porto Alegre: Editora da UFRGS, 2009. (Série Educação à Distância).

GIL, A. C. **Como elaborar projetos de pesquisa**. São Paulo: Atlas: 2002.

GILL, T. Metadata and the Web. In: BACA, M. (Ed.). **Introduction to metadata**. 3. ed. Los Angeles: Getty Research Institute, 2016.

GILLILAND, A. J. Setting the Stage. In: BACA, M. (Ed.). **Introduction to metadata**. 3. ed. Los Angeles: Getty Research Institute, 2016.

GLUSHKO, R. J. **The discipline of organizing**. 1 ed. Massachusetts, EUA: The MIT Press, 2013.

GOMES, A. L. A. Por uma semântica do patrimônio cultural. **Revista Ibero-Americana de Ciência da Informação**, Brasília, v. 9, n. 2, p. 441–459, 2016.

GOOGLE CULTURAL INSTITUTE. **About the Google Cultural Institute**. 2015.  
Disponível em: <<https://www.google.com/culturalinstitute/about>>. Acesso em: 8 jun. 2018.

GRANT, A. **CIMI XML Schema for SPECTRUM: Supporting Documentation**. CIMI Consortium, 2002. Disponível em: <<http://xml.coverpages.org/CIMI-SupportingDoc15.pdf>>. Acesso em: 23 mai. 2018.

GRUPO DE PESQUISA HUMANIDADES DIGITAIS. **Humanidades Digitais**, 2010.  
Disponível em: <<http://www.nehilp.org/~nehilp/HD/SobreHD.html>>. Acesso em: 8 jun. 2018.

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.

HYVÖNEN, E. Publishing and Using Cultural Heritage Linked Data on the Semantic Web. **Synthesis Lectures on the Semantic Web: Theory and Technology**, v. 2, n. 1, p. 1–159, 17 out. 2012.



ICOM/CIDOC DOCUMENTATION STANDARDS GROUP. **Definition of the CIDOC Conceptual Reference Model**: version 5.1.2, 2013. Disponível em: <[http://www.cidoc-crm.org/sites/default/files/cidoc\\_crm\\_version\\_5.1.2.pdf](http://www.cidoc-crm.org/sites/default/files/cidoc_crm_version_5.1.2.pdf)>. Acesso em: 21 mai. 2018.

ICOM INTERNATIONAL COMMITTEE FOR DOCUMENTATION. **What is LIDO**. 2010. Disponível em: <<http://network.icom.museum/cidoc/arbeitsgruppen/lido/L/6/>>. Acesso em: 03 abr. 2018.

ICONCLASS. 2012. Disponível em: <<http://www.iconclass.nl/home>>. Acesso em: 04 set. 2017.

ICONCLASS. **ICONCLASS as Linked Open Data**. 2015. Disponível em: <<http://www.iconclass.org/help/lod>>. Acesso em: 03 abr. 2018.

INSTITUTO BRASILEIRO DE MUSEUS. Tainacan: museus cariocas preparam-se para testar solução para acervos digitais. [201-]. Disponível em: <<http://www.museus.gov.br/tag/projeto-tainacan/>>. Acesso em: 11 set. 2018.

INSTITUTO DO PATRIMÔNIO HISTÓRICO E ARTÍSTICO NACIONAL (IPHAN). Patrimônio Mundial: fundamentos para seu reconhecimento. Brasília: IPHAN, 2008. Disponível em: <[http://portal.iphan.gov.br/uploads/ckfinder/arquivos/Cartilha\\_do\\_patrimonio\\_mundial.pdf](http://portal.iphan.gov.br/uploads/ckfinder/arquivos/Cartilha_do_patrimonio_mundial.pdf)>. Acesso em: 28 set. 2017.

INTERNATIONAL COUNCIL OF MUSEUMS; ICOM INTERNATIONAL COMMITTEE FOR DOCUMENTATION. What is the CIDOC CRM? [201-]. Disponível em: <<http://www.cidoc-crm.org/>>. Acesso em: 21 mai. 2018.

ISOTANI, S.; BITTENCOURT, I. I. **Dados Abertos Conectados**: em busca da Web do Conhecimento. 2015. Disponível em: <<http://ceweb.br/livros/dados-abertos-conectados/>>. Acesso em: 08 jun. 2017.

KIM, J. G.; HAUSENBLAS, M. **Five star open data**, 2015. Disponível em: <<http://5stardata.info/en/>>. Acesso em: 26 set. 2017.

KIM, S. et al. **Towards a semantic data infrastructure for heterogeneous Cultural Heritage data - challenges of Korean Cultural Heritage Data Model (KCHDM)**. IEEE, set. 2015. Disponível em: <<http://ieeexplore.ieee.org/document/7419508/>>. Acesso em: 29 jan. 2018.

KOLTAY, T. Library and information science and the digital humanities: perceived and real strengths and weaknesses. **Journal of documentation**, *online*, v. 72, n. 4, p. 781–792, 2016.

LEE, C. *et al.* Open Data meets Digital Curation: an investigation of practices and needs. **International Journal of Digital Curation**, *online*, v. 11, n. 2, p. 115–125, jul. 2017.

LIBRARY OF CONGRESS. **An introduction to VRA Core**. 2014. Disponível em: <[https://www.loc.gov/standards/vracore/VRA\\_Core4\\_Intro.pdf](https://www.loc.gov/standards/vracore/VRA_Core4_Intro.pdf)>. Acesso em: 03 abr. 2018.

LIBRARY OF CONGRESS. **Controlled Vocabularies**. [2017]. Disponível em: <<https://www.loc.gov/librarians/controlled-vocabularies/>>. Acesso em: 23 maio. 2018.

LIBRARY OF CONGRESS. **LC Linked Data Service: about**. [20--]. Disponível em: <<http://id.loc.gov/about/>>. Acesso em: 22 mai. 2018.

LIBRARY OF CONGRESS. **Standards at the Library of Congress**. 2015. Disponível em: <<http://www.loc.gov/standards/>>. Acesso em: 04 set. 2017.

LIBRARY OF CONGRESS. **Theaurus for Graphic Materials**. [201-]. Disponível em: <<http://id.loc.gov/vocabulary/graphicMaterials.html>>. Acesso em: 03 abr. 2018.

LINKED OPEN VOCABULARIES (LOV). **In LOV at a glance...** [201-]. Disponível em: <<https://lov.okfn.org/dataset/lov/about>>. Acesso em: 25 mai. 2018.

LINKED OPEN VOCABULARIES (LOV). **Metadata Recommendations For Linked Open Data Vocabularies**. 2012. Disponível em: <[https://lov.okfn.org/Recommendations\\_Vocabulary\\_Design.pdf](https://lov.okfn.org/Recommendations_Vocabulary_Design.pdf)>. Acesso em: 25 mai. 2018.

LIU, A. The state of the digital humanities: A report and a critique. **Arts and Humanities in Higher Education**, v. 11, n. 1–2, p. 8–41, 2012.

LODLAM.NET. **About**. 2010. Disponível em: <<http://lodlam.net/about/>>. Acesso em: 8 jun. 2018.

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

MARCONDES, C. H. Linked data - dados interligados - e interoperabilidade entre arquivos, bibliotecas e museus na web. **Encontros Bibli: revista eletrônica de biblioteconomia e ciência da informação**, Florianópolis, v. 17, n. 34, 9 ago. 2012.

MAROEVIC, I. The phenomenon of cultural heritage and the definition of a unit of material. **Nordisk Museologi**, *online*, v. 2, 1998, p.135–142. Disponível em: <<https://www.journals.uio.no/index.php/museolog/article/view/3829>>. Acesso em: 14 mar. 2017.

MARON, N. L.; YUN, J.; PICKLE, S. **Sustaining our digital future**: institutional strategies for digital content. New York: Ithaca S+R, 2013.

MCCRAE, J. P. *et al.* **The Linked Open Data Cloud**. 2018. Disponível em: <<http://lod-cloud.net/>>. Acesso em: 8 jun. 2018.

MÉNDEZ RODRÍGUEZ, E. M.; GREENBERG, J. Linked data for open vocabularies and HIVE's global framework. 2012.

MITCHELL, E. T. Three Case Studies in Linked Open Data. **Library Technology Reports**, [s.l.], v. 49, n. 5, p. 26-43, jul. 2013. American Library Association. Disponível em: <<https://journals.ala.org/index.php/ltr/article/view/4693>>. Acesso em: 09 jun. 2018.

MIXTER, J. *et al.* **VRA Ontology**: vocabulary for representing entities in the VRA data model. [2015]. Disponível em: <<https://s3.amazonaws.com/VRA/ontology.html#>>. Acesso em: 21 mai. 2018.

NININ, D. M.; SIMIONATO, A. C. Linked Open Data e sustentabilidade de acervos digitais de patrimônios culturais. **IRIS: Revista de Informação, Memória e Tecnologia**, Recife, v. 3, n. especial, p. 93-110, 2014/2017. Disponível em: <<https://periodicos.ufpe.br/revistas/IRIS/article/view/236190/29019>>. Acesso em: 09 mai. 2018.

NOWACK, B. **The Semantic Web Technology Stack (not a piece of cake...)**, 2009. Disponível em: <<https://www.flickr.com/photos/jalbertbowdenii/15136351742/>>. Acesso em: 8 jun. 2018.

ONLINE COMPUTER LIBRARY CENTER (OCLC). **CIMI Consortium**: A Valued Association. 2018. Disponível em: <<https://www.oclc.org/research/activities/cimi.html>>. Acesso em: 23 mai. 2018.

ONLINE COMPUTER LIBRARY CENTER (OCLC). **OCLC Linked Data**. EUA: OCLC, 2015. Disponível em: <<https://www.oclc.org/developer/develop/linked-data.en.html>>. Acesso em: 13 abr. 2017.

OPEN ARCHIVES INITIATIVE. **Object Exchange and Reuse**. [2008]?. Disponível em: <<https://www.openarchives.org/ore/>>. Acesso em: 8 jun. 2018.

OPEN ARCHIVES INITIATIVE. **Open Archives Initiative Protocol for Metadata Harvesting**. [199-]. Disponível em: <<https://www.openarchives.org/pmh/>>. Acesso em: 04 set. 2017.

OPEN KNOWLEDGE. **Open Data Commons**. [201-]. Disponível em: <<https://opendatacommons.org/licenses/>>. Acesso em: 8 jun. 2018.

OPEN KNOWLEDGE. **OpenGlam**. [201-]. Disponível em: < <https://openglam.org/>>. Acesso em: 21. jun. 2016.

ORGANIZAÇÃO DAS NAÇÕES UNIDAS PARA A EDUCAÇÃO, A CIÊNCIA E CULTURA (UNESCO). Convenção para a Proteção do Patrimônio Mundial, Cultural e Natural. Unesco, Paris, 1972. In: INSTITUTO DO PATRIMÔNIO HISTÓRICO E ARTÍSTICO NACIONAL (IPHAN). **Patrimônio Mundial: fundamentos para seu reconhecimento**. Brasília: IPHAN, 2008. p. 58-80. Disponível em: <[http://portal.iphan.gov.br/uploads/ckfinder/arquivos/Cartilha\\_do\\_patrimonio\\_mundial.pdf](http://portal.iphan.gov.br/uploads/ckfinder/arquivos/Cartilha_do_patrimonio_mundial.pdf)>. Acesso em: 28 set. 2017.

ORGANIZAÇÃO DAS NAÇÕES UNIDAS PARA A EDUCAÇÃO, A CIÊNCIA E CULTURA (UNESCO). **Convenção para a Salvaguarda do Patrimônio Cultural Imaterial**. Tradução do Ministério das Relações Exteriores. Brasília, 2006. Disponível em: <<http://portal.iphan.gov.br/uploads/ckfinder/arquivos/ConvencaoSalvaguarda.pdf>>. Acesso em: 19 jul. 2017.

ORGANIZAÇÃO DAS NAÇÕES UNIDAS PARA A EDUCAÇÃO, A CIÊNCIA E CULTURA (UNESCO). **Declaração de Budapeste sobre o Patrimônio Mundial**. 2002. Disponível em: <<http://www.patrimoniocultural.gov.pt/media/uploads/cc/declaracaoBudapestesobrepatrimoniomundial2002.pdf>>. Acesso em: 11 jul. 2017.

ORGANIZAÇÃO DAS NAÇÕES UNIDAS PARA A EDUCAÇÃO, A CIÊNCIA E CULTURA (UNESCO). **Manifesto para a Preservação Digital**. Disponível em: <[http://www.bnportugal.pt/agenda/ecpa/manifesto\\_unesco.html](http://www.bnportugal.pt/agenda/ecpa/manifesto_unesco.html)>. Acesso em: 8 jun. 2018.

POMERANTZ, J. **Metadata**. Cambridge, Massachusetts ; London, England: The MIT Press, 2015.

PORSDAM, H. Digital Humanities: On Finding the Proper Balance between Qualitative and Quantitative Ways of Doing Research in the Humanities. **DHQ: Digital Humanities Quarterly**, *online*, v. 7, n. 3, 2013.

RAMALHO, R. A. S. **Web Semântica: aspectos interdisciplinares da gestão de recursos informacionais no âmbito da Ciência da Informação**. 2006. 120 f. Dissertação (mestrado) - Universidade Estadual Paulista, Faculdade de Filosofia e Ciências, 2006. Disponível em: <<http://repositorio.unesp.br/handle/11449/93709>>. Acesso em: 07 abr. 2016.

RAMALHO, R. A. S.; VIDOTTI, S. A. B. G.; FUJITA, M. S. L. Web semântica: uma investigação sob o olhar da Ciência da Informação. **DataGramaZero**, Rio de Janeiro, v. 8, n. 6, dez/2007.

RASMUSSEN-PENNINGTON, D. Demystifying Linked Data: are you ready for what's next? CILIP Update, online, jul./ago, 2016.

RAY, E. T. **Aprendendo XML**. Rio de Janeiro: Campus, 2001.

REVEZ, J. Práticas digitais nas unidades de I&D em Portugal: uma observação parcial da área das Humanidades. **Informação & Informação**, Londrina, v. 22, n. 3, p. 405–425, 2017.

RILEY, J. **Understanding Metadata: what is metadata, and what is it for?** National Information Standards Organization (NISO), , 2017. Disponível em: <[http://www.niso.org/apps/group\\_public/download.php/17446/Understanding%20Metadata.pdf](http://www.niso.org/apps/group_public/download.php/17446/Understanding%20Metadata.pdf)>. Acesso em: 01 ago. 2017.

ROMANETTO, L. M.; SANTOS, P. L. V. A. C.; ALVES, R. C. V. O Virtual International Authority File - VIAF e a agregação de valores por metadados de autoridade. **RDBCI: Revista Digital Biblioteconomia e Ciência da Informação**, Campinas, v. 15, n.3, set./dez. 2017.

ROMANETTO, L. M. **O controle de autoridade no consórcio VIAF**. 2017.104 f. Dissertação (Mestrado em Ciência da Informação). Universidade Estadual Paulista Júlio de Mesquita Filho, Faculdade de Filosofia e Ciências de Marília, 2017. Disponível em: <<https://repositorio.unesp.br/handle/11449/148839>>. Acesso em: 22 fev. 2018.

SANTAREM SEGUNDO, J. E. Tecnologías de la información y la comunicación para proporcionar datos abiertos en formato semántico. **Ibersid: revista de sistemas de información y documentación**, *online*, v. 7, p. 33–40, 2013. Disponível em: <<http://ibersid.eu/ojs/index.php/ibersid/article/view/4075/3744>>. Acesso em: 04 out. 2017.

SANTAREM SEGUNDO, J. E. Web semântica, dados ligados e dados abertos: uma visão dos desafios do brasil frente às iniciativas. **Tendências da Pesquisa Brasileira em Ciência da Informação**, João Pessoa, v. 8, n. 2, 2015.

SANTAREM SEGUNDO, J. E.; SIMIONATO, A. C. **Uma abordagem sobre a estrutura do Geonames e suas contribuições para o Linking Open Data**. Anais do XVII Encontro Nacional de Pesquisa em Ciência da Informação. **Anais...** In: XVII ENANCIB. Salvador, BA: UFBA, 2016.

SANTOS NETO, A. L. *et al.* Tecnologias de dados abertos para interligar bibliotecas, arquivos e museus: um caso machadiano. **Transinformação**, Campinas, v. 25, n. 1, 2013.

SANTOS, H. P. Modelo CIDOC CRM: interoperabilidade semântica de informações culturais. **Brazilian Journal of Information Science**, *online*, v. 10, n. 1, 2016. Disponível em: <<http://www.brapci.ufpr.br/brapci/v/a/19615>>. Acesso em: 22 fev. 2018.

SANTOS, P. L. V. A. C.; ALVES, R. C. V. Metadados e Web Semântica para estruturação da Web 2.0 e Web 3.0. **DataGramZero**, Rio de Janeiro, v. 6, n. 10, 2009. Disponível em: <<http://www.brapci.ufpr.br/brapci/index.php/article/download/52958>>. Acesso em: 02 ago. 2017.

SANTOS, P. L. V. A. C.; SANTANA, R. C. G. Dado e Granularidade na perspectiva da Informação e Tecnologia: uma interpretação pela Ciência da Informação. **Ciência da Informação**, Brasília, v. 42, n. 2, p. 199–209, 2013.

SANTOS, P. L. V. A. DA C.; SIMIONATO, A. C.; ARAKAKI, F. A. Definição de metadados para recursos informacionais: apresentação da metodologia BEAM. **Informação & Informação**, Londrina, v. 19, n. 1, p. 146, 25 fev. 2014.

SARACEVIC, T. Ciência da Informação: origem, evolução e relações. Tradução: Ana Maria Pereira Cardoso. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 1, n. 1, p. 41-62, jan./jun. 1996.

SARACEVIC, T. Interdisciplinary nature of information science. **Ciência da informação**, v. 24, n. 1, p. 36–41, 1995.

SILVA, J. L. C.; FREIRE, G. H. A. Um olhar sobre a origem da ciência da informação: indícios embrionários para sua caracterização identitária. *Encontros Bibli: revista eletrônica de biblioteconomia e ciência da informação*, v. 17, n. 33, p. 1-29, jan./abr. 2012.

SIMIONATO, A. C.; ARAKAKI, F. A.; SANTOS, P. L. V. A. C. Descrição em bibliotecas, arquivos, museus e galerias de arte: linkando recursos e comunidades. **Informação & Informação**, Londrina, v. 22, n. 2, p. 449, 29 out. 2017.

SMITH, L. **Uses of heritage**. London; New York: Routledge, 2008

SMITH, M. **Proposed a 4-star classification-scheme for linked open cultural metadata**, 2011. Disponível em: <<http://lodlam.net/2011/06/06/proposed-a-4-star-classification-scheme-for-linked-open-cultural-metadata/>>. Acesso em 21 de mar. 2017.

SOCIEDADE DE USUÁRIOS DE COMPUTADORES E EQUIPAMENTOS SUBSIDIÁRIOS (SUCE SU). **Dicionário de termos técnicos em computação: inglês/português**. [19--]. 72p.

SOUZA, W. E. R.; CRIPPA, G. O campo da ciência da informação e o patrimônio cultural: reflexões iniciais para novas discussões sobre os limites da área. **Encontros Bibli: revista eletrônica de biblioteconomia e ciência da informação**, Florianópolis, v. 15, n. 29, 25 maio 2010.

SOUZA, W. E. R.; CRIPPA, G. O patrimônio como processo: uma idéia que supera a oposição material-imaterial. **Em Questão**, Porto Alegre, v. 17, n. 2, p. 241-255, jul/dez. 2011. Disponível em: <<http://www.brapci.ufpr.br/brapci/v/a/11525>>. Acesso em: 13 abr. 2017.

SPIRO, L. This is why we fight: defining the values of the digital humanities. **Debates in the digital humanities**, 16 p., 2012.

SUPPLEJACK. **About**. 2013. Disponível em: <<https://digitalnz.github.io/supplejack/about.html>>. Acesso em: 29 mai. 2018.

THE GETTY RESEARCH INSTITUTE. **Categories for Description of Works of Art**. 2017. Disponível em: <[http://www.getty.edu/research/publications/electronic\\_publications/cdwa/index.html](http://www.getty.edu/research/publications/electronic_publications/cdwa/index.html)>. Acesso em: 22 mai. 2018.

THE GETTY RESEARCH INSTITUTE. **Categories for Description of Works of Art: describe, catalogue, link works of art, architecture, and cultural heritage**. 2016. Disponível em: <[http://www.getty.edu/research/tools/vocabularies/vocab\\_cdwa\\_flier.pdf](http://www.getty.edu/research/tools/vocabularies/vocab_cdwa_flier.pdf)>. Acesso em: 29 ago. 2017.

THE GETTY RESEARCH INSTITUTE. **CDWA Lite**. 2013. Disponível em: <[https://www.getty.edu/research/publications/electronic\\_publications/cdwa/cdwalite.html](https://www.getty.edu/research/publications/electronic_publications/cdwa/cdwalite.html)>. Acesso em: 03 abr. 2018.

THE GETTY RESEARCH INSTITUTE. **Getty Vocabularies**. [20--]. Disponível em: <<http://www.getty.edu/research/tools/vocabularies/>>. Acesso em: 23 maio. 2018.

THE GETTY RESEARCH INSTITUTE. **Getty vocabularies as Linked Open Data**. 2017. Disponível em: <<https://www.getty.edu/research/tools/vocabularies/lod/index.html>>. Acesso em: 03 abr. 2018.

THORSEN, H. K.; PATTUELLI, M. Cristina. Linked Open Data and the cultural heritage landscape. In: JONES, Ed; SEIKEL, Michele (Eds.). **Linked data for cultural heritage**. Chicago: ALA Editions, 2016. p. 1-22.

UNIÃO EUROPEIA. **European Comission**. 2017. Disponível em: <[https://europa.eu/european-union/about-eu/institutionsbodies/europeancommission\\_en](https://europa.eu/european-union/about-eu/institutionsbodies/europeancommission_en)>. Acesso em: 02 out. 2017.

VAN HOOLAND, S.; VERBORGH, R. **Linked Data for Libraries, Archives and Museums: How to clean, link and publish your metadata**. [s.l.]: Facet publishing, 2014.  
VANDEGRIFT, M. What is digital humanities and what's it doing in the library. **The Library with the Lead Pipe**, *online*, 27 junho de 2012. Disponível em: <<http://www.inthelibrarywiththeleadpipe.org/2012/dhandthelib/>>. Acesso em: 05 jul. 2017.

VISUAL RESOURCES ASSOCIATION. **Cataloging Cultural Objects**. 2006. Disponível em: <<http://cco.vrafoundation.org/index.php/aboutindex/>>. Acesso em: 23 maio. 2018.

VISUAL RESOURCES ASSOCIATION. **Metadata Standards of Cataloging Cultural Objects**. 2006. Disponível em: <[http://cco.vrafoundation.org/index.php/toolkit/metadata\\_standards/](http://cco.vrafoundation.org/index.php/toolkit/metadata_standards/)>. Acesso em: 8 jun. 2018.

VISUAL RESOURCES ASSOCIATION. **VRA Core RDF Ontology Available for Review**, 2015. Disponível em: <<http://vraweb.org/vra-core-rdf-ontology-available-for-review/>>. Acesso em: 8 jun. 2018

VOSS, J. **Radically open cultural heritage data on the Web**. 2012. Disponível em: <[http://www.museumsandtheweb.com/mw2012/papers/radically\\_open\\_cultural\\_heritage\\_data\\_on\\_the\\_w](http://www.museumsandtheweb.com/mw2012/papers/radically_open_cultural_heritage_data_on_the_w)>. Acesso em: 31 maio. 2017.

WINER, D.; ROCHA, I. E. Europeana: um projeto de digitalização e democratização do patrimônio cultural europeu. **Patrimônio e Memória**, Assis, v. 9, n. 1, p. 113–127, 2013.

WOOD, D. *et al.* **Linked data**: structured data on the Web. Shelter Island, NY: Manning, 2014.

WORLD WIDE WEB CONSORTIUM. **Data on the Web Best Practices**. 2017. Disponível em: <<https://www.w3.org/TR/dwbp/>>. Acesso em: 23 maio. 2018.

WORLD WIDE WEB CONSORTIUM. **Extensible Markup Language (XML)**. 2016. Disponível em: <<https://www.w3.org/XML/>>. Acesso em: 23 maio. 2018.

WORLD WIDE WEB CONSORTIUM. **Linked data**. 2015. Disponível em: <<https://www.w3.org/standards/semanticweb/data>>. Acesso em: 11 jun. 2018.

WORLD WIDE WEB CONSORTIUM. **Resource Description Framework (RDF)**. 2014. Disponível em: <<https://www.w3.org/2001/sw/wiki/RDF>>. Acesso em: 11 jun. 2018b.

WORLD WIDE WEB CONSORTIUM. **Semantic Web Education and Outreach (SWEO) Interest Group**. 2008. Disponível em: <<https://www.w3.org/blog/SWEO/>>. Acesso em: 8 jun. 2018.