

Ontologias para anotação semântica em mídias: uma construção colaborativa de redes de conhecimento do patrimônio cultural

Daniela Lucas da Silva Lemos¹

Abeil Coelho-Júnior²

Danielle do Carmo³

Resumo: o artigo objetiva apresentar um levantamento sistemático de modelos ontológicos que estão sendo comumente usados como padrões de anotação em mídias diversas, sobretudo endereçados a projetos de integração de acervos culturais em ambiente *Linked Open Data* e envolvendo as plataformas da Fundação Wikimedia conhecidas como Wikipédia, Wikimedia Commons e Wikidata, em que se busca ampliar a base de usuários e a produção colaborativa de seus recursos de informação visando ampliar a difusão das coleções patrimoniais, tornar explícito o conhecimento gerado em seus acervos, e, logo, valorizar e democratizar a cultura na sociedade em rede. Trata-se de uma pesquisa teórica e qualitativa, de cunho exploratório e descritivo usando como método uma revisão sistemática de literatura em bases de dados especializadas nos campos das Ciências da Informação e da Computação juntamente com o uso de dimensões analíticas para organizar elementos constitutivos das ontologias elencadas no estudo. Como resultados, apresenta uma análise sistemática dos modelos conceituais ontológicos CIDOC CRM, CRMdig, Linked Art, EMD e FRBRoo, mostrando como esses padrões de documentação estão sendo usados como formas de representação do conhecimento multimídia no contexto do patrimônio cultural e como organizam em suas estruturas, metadados de natureza multimídia (independentes, dependentes e descritivos de conteúdo) e de proveniência (origem e métodos de produção) para descrição de objetos culturais em variados contextos e necessidades. Os modelos analisados apresentam-se como bons candidatos a reúso, sobretudo em projetos que visem a integração de acervos do patrimônio cultural em ambiente *Linked Open Data* e envolvendo o ecossistema wiki.

Palavras-chave: anotação semântica; ontologias; patrimônio cultural; multimídia; fundação wikimedia.

¹ Doutora em Ciência da Informação pela Universidade Federal de Minas Gerais (UFMG); mestre em Ciência da Informação pela Universidade Federal de Minas Gerais (UFMG); danielalucas@hotmail.com. Lattes (opcional). ORCID (opcional).

² Mestrando em Ciência da Informação pela UFES, bacharel em Arquivologia pela Universidade Federal do Espírito Santo (UFES) abeilm@hotmail.com. ORCID (opcional).

³ Doutoranda em Ciência da Informação pela Universidade de Brasília (UnB). Mestre em Memória Social e Patrimônio Cultural pela Universidade Federal de Pelotas (UFPEL). docarmo.danielle@gmail.com. Lattes (opcional). ORCID (opcional).

ONTOLOGIES FOR SEMANTIC ANNOTATION IN MEDIA: A COLLABORATIVE CONSTRUCTION OF CULTURAL HERITAGE KNOWLEDGE NETWORKS

Abstract: this article aims to present a systematic survey of ontological models that are being commonly used as annotation standards in diverse media, especially addressed to projects of integration of cultural collections in a connected open data environment and involving as platforms of the Wikimedia Foundation such as Wikipedia, Wikimedia Commons and Wikidata, which seeks to expand the user base and the collaborative production of its information resources to expand the dissemination of heritage collections, make explicit what is generated in its known collections, and, therefore, value and democratize culture in the network society. This is theoretical and qualitative research, exploratory and descriptive, using as a method a systematic literature review in specialized databases in the fields of Information and Computer Science, together with the use of analytical dimensions to organize constituent elements of ontologies listed in the study. As a result, it presents a systematic analysis of CIDOC CRM, CRMdig, Linked Art, EMD and FRBRoo conceptual ontology models, showing how these document standards are being used as forms of representation of multimedia knowledge in the context of cultural heritage and how they organize in its structures, multimedia metadata (independent, dependent and descriptive of content) and provenance (origin and production methods) for the description of cultural objects in different contexts and needs. The models are presented as good candidates for reuse, especially in projects aimed at integrating cultural heritage collections in a Linked Open Data environment and involving the wiki's ecosystem.

Keywords: semantic annotation; ontologies; cultural heritage; multimedia; wikimedia foundation.



1 INTRODUÇÃO

Nos últimos anos, variados setores, incluindo saúde, negócio, entretenimento, ciência, governo, educação e cultura (DOMINGUE; FENSEL; HENDLER, 2011; MORA-CANTALLOPS; SÁNCHEZ-ALONSO; GARCÍA-BARRIOCANAL, 2019) se veem imersos pelas inovações proporcionadas pelas mídias digitais integradas na internet, cujo crescimento e disponibilidade são fortemente impulsionados por dispositivos móveis, como tablets, smartphones, câmeras digitais, entre outros, com os quais se potencializa a produção e o consumo de conteúdo multimídia na rede.

No contexto do setor cultural, foco desta pesquisa, instituições ligadas ao campo do patrimônio cultural vêm se interessando em tornar mais efetivo o acesso e a recuperação documental por meio da aplicação de tecnologias que viabilizam o tratamento de dados na web (DIJKSHOORN et al., 2018; EUROPEANA TECH, 2020; ROLEDANO-ARILLO; NAVARRO-BONILLA; CERDÁ-DÍAZ, 2019; TRIQUES; ARAKAKI, 2021; PESCHANSKI, 2021), ambiente no qual, atualmente, se manifesta grande parte de seus acervos, a exemplo do que ocorre nas denominadas GLAM (do inglês, *Galleries, Libraries, Archives, Museums*) de “galerias, bibliotecas, arquivos e museus”, através das quais se mantêm grandes repositórios de textos, imagens, vídeos, sons, modelos 3D, *games*, enfim, recursos multimídia diversos.

A digitalização dos acervos analógicos destas instituições é a chave para que se tornem parte integrante da web e, desse modo, consigam potencializar a democratização do conhecimento científico e cultural à sociedade em rede. Adicionalmente, uma vez digitais e parte de uma rede aberta, os artefatos culturais podem ser compartilhados, recomendados, remixados, misturados, incorporados e citados (OOMEN; AROYO, 2011), valorizando a colaboração e o reúso de dados por seus usuários.

Uma iniciativa que vem se destacando no contexto GLAM é a possibilidade de integração de seus acervos culturais com as plataformas da Fundação Wikimedia (WIKIMEDIA FOUNDATION, 2021) conhecidas como Wikipédia (WIKIPÉDIA, 2021), Wikimedia Commons (WIKIMEDIA COMMONS, 2021) e Wikidata (WIKIDATA, 2021), em que se busca ampliar a base de usuários e a produção colaborativa de seus recursos de informação visando a participação em redes de conhecimento para interesses diversos.

A Wikipédia, lançada em 2001, é uma enciclopédia *online*, que tem o seu modo de produção de conteúdo baseado na construção colaborativa e curadoria coletiva, assim como as demais



plataformas geridas pela Wikimedia Foundation. Segundo informações fornecidas pela plataforma, a Wikipédia “tem como propósito fornecer um conteúdo livre, objetivo e verificável, que todos possam editar e melhorar” (WIKIPÉDIA, 2021). A plataforma tem uma proposta multilíngue, ou seja, existem projetos de Wikipédias em 320 idiomas. Os editores da Wikipédia, conhecidos como wikipedistas, são voluntários que atuam de forma colaborativa, seja por meio de esforços individuais ou coordenados coletivamente. Para se ter uma noção das dimensões da Wikipédia, somente o domínio pt.wikipedia.org possui atualmente 1.079.205 páginas de conteúdo em português, 2.620.327 editores registrados e 9.556 editores ativos, ou seja, que realizaram alguma ação nos últimos 30 dias⁴.

O Wikimedia Commons é um repositório de mídias multilíngue que alimenta todos os outros projetos da Wikimedia Foundation. A plataforma tem como objetivo distribuir conteúdo educacional de forma livre e gratuita, ou seja, as licenças devem permitir o uso, a cópia e a modificação dos arquivos disponibilizados, respeitando os termos especificados pelos autores (WIKIMEDIA COMMONS, 2021). O repositório de mídias Wikimedia Commons foi inaugurado em 2004 e atualmente apresenta um acervo de 75.110.434 arquivos de mídias, 10.295.602 usuários registrados e 37.686 editores ativos⁵.

Por sua vez, o Wikidata é apresentado como “uma base de conhecimento livre e aberta que pode ser lida e editada por humanos e máquinas” (WIKIDATA, 2021). Desta forma, a plataforma funciona como um repositório de dados e declarações estruturados que além de alimentar os projetos irmãos, também alimenta serviços como o Google e a assistente de voz Siri (VILLAESPESA; NAVARRETE, 2019). Isso é possível uma vez que os conteúdos estão disponíveis sob uma licença aberta que permite a livre reutilização. A Wikidata foi lançada em 2012 e atualmente apresenta 94.278.527 páginas de itens, 5.031.248 usuários registrados e 22.871 usuários ativos⁶.

No continente europeu, destacam-se projetos e oportunidades de colaboração entre comunidades do patrimônio cultural e do Wikidata, sendo esta usada como base de conhecimento de referência (DAQUINO et al., 2017; MORA-CANTALLOPS; SÁNCHEZ-ALONSO; GARCÍA-BARRIOCANAL, 2019; METILLI; BARTALESI; MEGHINI, 2019) na descrição de entidades à luz de declarações semânticas disponíveis na plataforma, o que permite gerar

⁴ Dados coletados em 23/07/2021 em: <https://pt.wikipedia.org/wiki/Especial:Estatísticas>

⁵ Dados coletados em 23/07/2021 em: <https://commons.wikimedia.org/wiki/Special:Statistics?uselang=pt-br>

⁶ Dados coletados em 23/07/2021 em: <https://www.wikidata.org/wiki/Special:Statistics>

gráficos de conhecimento capazes de agregar semântica, por exemplo, em entidades presentes nas páginas do Wikipédia, e em arquivos publicados no Wikimedia Commons por meio de descrições em seus aspectos de mídia, de conteúdo e de proveniência, potencializando busca, recuperação e reúso.

No Brasil, nota-se uma carência de iniciativas nessa perspectiva, muito em função do emprego precário de soluções sociotécnicas contemporâneas (NÚCLEO DE INFORMAÇÃO E COORDENAÇÃO DO PONTO BR, 2019) endereçadas à organização, à preservação, ao acesso e à recuperação da informação na rede. Entretanto, torna-se oportuno mencionar a presença de acervos de museus ligados ao Instituto Brasileiro de Museus (Ibram) no repositório Wikimedia Commons com imagens relacionadas a pinturas (CARMO; MARTINS, 2020), além da existência de páginas de museus ligados à mesma instituição na rede da enciclopédia Wikipédia. Iniciativa semelhante se apresenta no Arquivo Nacional, localizado na cidade do Rio de Janeiro, no museu de Anatomia Veterinária e no museu Paulista, pertencentes à Universidade de São Paulo, em que se tem registro de várias imagens e artigos publicados também no repositório de mídias e na enciclopédia wiki, com algumas atuações de integração com o Wikidata (PESCHANSKI, 2021).

Logo, percebe-se a falta de iniciativas mais encorajadas por parte das instituições culturais brasileiras voltadas à estratégia de publicação e integração de seus acervos com entidades do Wikidata, o que poderia ampliar buscas e recuperação de informações culturais no Wikimedia Commons e nas páginas na rede do Wikipédia, possibilitando a construção colaborativa de redes de conhecimento do patrimônio cultural a partir do ecossistema wiki, além de bibliotecas digitais externas, como, por exemplo, a Europeana (CHARLES; CLAYPHAN; ISAAC, 2017).

Um dos desafios levantados a partir desse contexto GLAM-Wiki, se refere à integração de dados institucionais dos acervos digitais e disponíveis em plataformas abertas e as bases de dados heterogêneas disponíveis no ecossistema wiki, o que demanda requisitos de interoperabilidade a nível sintático e semântico pelo fato de esses dados serem usados em diferentes comunidades. No nível sintático, os dados precisam ser representados em estruturas e formatos similares; já no nível semântico, os esquemas de metadados e vocabulários associados precisam ter significado livre de ambiguidade (ZENG, 2019) presente na linguagem natural. Interoperabilidade, portanto, consiste na capacidade de diversos sistemas trocarem dados entre si (interoperar) com perda mínima de conteúdo e funcionalidade, de modo a garantir que pessoas, organizações e sistemas computacionais interajam satisfatoriamente (RILEY, 2017).



Encaminhamentos de pesquisas têm sido feitos progressivamente pela Ciência da Informação (CI) e Ciência da Computação (CC) cujos esforços são endereçados em melhorias aos sistemas de informação, especialmente aos sistemas de recuperação da informação, sobretudo no que se refere à organização e à representação do conhecimento sobre os recursos de informação envolvidos. Podem-se citar a Web Semântica (BERNERS-LEE; HENDLER; LASSILA, 2001) e sua proposta emergente de dados abertos ligados ou *Linked Open Data* (LOD) (BIZER; HEATH; BERNERS-LEE, 2009; MACHADO; SOUZA; SIMÕES, 2019); instrumentos de representação de relacionamentos semânticos e conceituais como ontologias e vocabulários controlados (GUIZZARDI, 2005; ALMEIDA, 2013; LEMOS; SOUZA, 2020); e modelos conceituais de referência que orientam a modelagem da realidade documental e o processo de busca e recuperação da informação em contextos digitais como o FRBR – *Functional Requirements for Bibliographic Records* (IFLA, 2009) e o CIDOC CRM – *International Committee for Documentation/Conceptual Reference Model* (BOEUF et al., 2018).

Nessa perspectiva, torna-se válido ressaltar que a abordagem técnica para a integração de acervos culturais GLAM com o ecossistema wiki lança mão dos preceitos do modelo LOD por meio do qual se faz uso de padrões abertos para estabelecer interligações e anotações (a partir de *Uniform Resource Identifier* - URI) sobre dados sob licença aberta. A proposta possibilita que provedores de conteúdo multimídia (instituições de cultura, por exemplo) enriqueçam semanticamente seus esquemas de metadados com especificações estruturadas e bem definidas de conhecimento, tais como padrões de metadados, modelos conceituais, vocabulários controlados e ontologias, viabilizando o consumo e o reúso de informações de alta qualidade. Busca-se, portanto, uma visão integrada de dados e uma maximização da interoperabilidade entre conjuntos de dados (do inglês, *datasets*) de produtores e consumidores de acervos culturais multimídia na rede para fins de descoberta de novos conhecimentos.

Para o propósito da presente pesquisa, ontologias são consideradas modelos de anotação mais sofisticados (LEMOS; SOUZA, 2019; LEMOS, 2020; LEMOS; SOUZA, 2020) em termos de tratamento semântico da informação disponível em rede, o que permite um usuário (humano ou agente computacional) descrever e interligar semanticamente recursos multimídia por meio de qualificadores como conceitos, instâncias, propriedades, relações e restrições mantidas entre tais recursos. Tal abordagem é denominada na literatura de anotação semântica (UREN et al., 2006; SHADBOLT; BERNERS-LEE; HALL, 2006), cuja fundamentação subjaz aos preceitos da Web semântica e do LOD, e tem como propósito identificar formalmente conceitos e relações



entre conceitos em documentos, tornando-os mais inteligentes e facilitando, portanto, o processo de interpretação dos dados pelos sistemas de recuperação de informação.

Desse modo, a partir de uma carência de projetos observada na revisão de literatura, especialmente envolvendo instituições culturais brasileiras na publicação de seus acervos no Wikimedia Commons e no Wikidata, torna-se oportuno a proposição de um modelo conceitual orientado a ontologias que sirva de base para a agregação de acervos culturais brasileiros nas redes de informação das plataformas wiki. Para tal, uma estratégia de modelagem adotada por grande parte das metodologias de construção de ontologias (SILVA; SOUZA; ALMEIDA, 2008; SUÁREZ-FIGUEROA; GÓMEZ-PÉREZ; FERNÁNDEZ-LÓPEZ, 2012; MENDONÇA; SOARES, 2017) está na inclusão de procedimentos adequados a seleção e a reúso de uma variedade de vocabulários e ontologias disponíveis e que precisam ser compatíveis, neste caso, com as entidades reais que organizam e representam o universo multimídia e contextual das bases de dados do Wikimedia, do Wikidata e dos acervos institucionais de cultura, de modo a conseguirem interoperar seus dados em rede.

Nesse viés, o objetivo do artigo é apresentar um levantamento sistemático de modelos ontológicos que estão sendo comumente usados como padrões de anotação em mídias diversas, sobretudo endereçados a projetos de integração de acervos culturais em ambiente LOD e envolvendo o ecossistema wiki. A partir desse rol de vocabulários semânticos candidatos a reúso que caracterize entidades (dados e metadados) de natureza multimídia, surgem oportunidades e desafios em responder às seguintes questões de pesquisa: *como os padrões de documentação orientados por ontologias estão sendo usados como formas de representação do conhecimento multimídia com o qual se realiza o contexto de um domínio do patrimônio cultural? e como esses vocabulários semânticos buscam organizar tipologias de metadados para descrição de conteúdo multimídia frente a variados contextos e necessidades?*

Espera-se que esta pesquisa sirva de base futura para projetos que demandem a implementação de ontologias nessa perspectiva e necessitem de um levantamento atual e consistente destinado a uma fase metodológica estratégica, a saber, de seleção e reúso de padrões adequados à anotação semântica de itens de coleções pertencentes a instituições de cultura por meio dos quais seus metadados e mídias associadas poderão ser publicados e integrados, respectivamente, no Wikidata e no Commons a partir de declarações semânticas constitutivas da base de conhecimento do Wikidata.



2 MODELOS DE ESTRUTURAÇÃO DA INFORMAÇÃO EM REDE: USO DE ANOTAÇÕES SEMÂNTICAS

Existem vários padrões de metadados, vocabulários controlados e ontologias existentes endereçados a estruturação da informação em rede para fins de integração de bases de dados oriundas de acervos do patrimônio cultural (FINK, 2018; DIJKSHOORN et al., 2018; EUROPEANA TECH, 2020). Os padrões de metadados têm a capacidade de prover um vocabulário comum para descrever uma variedade de estruturas de dados capazes de satisfazer a várias comunidades (ABBAS, 2010; GILLILAND, 2016). Como exemplo desses padrões têm-se o *Lightweight Information Describing Objects* (LIDO)⁷ para museus; o *Encoded Archival Description* (EAD)⁸ para arquivos; o *Machine Readable Cataloging Record* (MARC)⁹ para as bibliotecas; o *Visual Resources Association Core Categories* (VRA Core)¹⁰ para a descrição de obras de cultura visual, bem como das imagens que as representam; o Dublin Core¹¹ para a descrição de qualquer tipo de recurso, incluindo várias coleções de documentos e mídias; e o padrão MPEG-7¹² para prover um vocabulário multimídia rico e comum.

Apesar de esses padrões de metadados se encontrarem em formato de serialização padronizado pelo *World Wide Web Consortium* (W3C), como nas linguagens *eXtensible Markup Language* (XML), *Resource Description Framework* (RDF), *RDF Schema* ou mesmo *Ontology Web Language* (OWL) para representações de características de objetos digitais na web (ALLEMANG; HENDLER; GANDON, 2020), para o propósito desta pesquisa, tais padrões não são considerados ontologias pelo pouco rigor formal expressado em suas estruturas, geralmente compostas de classes facilmente compreensíveis e de relações comuns entre estas classes, não incluindo relações especiais, por exemplo, relações lógicas; e outros tipos de primitivas de representação, por exemplo, axiomas (GÓMEZ-PÉREZ, 2004; ALMEIDA, 2013).

Além de padrões de metadados, instituições de cultura inseridas em ambientes digitais vêm mantendo seus próprios vocabulários alinhados com sistemas de organização do conhecimento (SOCs) (HJØRLAND, 2003; 2007; 2015) menos formais, se comparados aos modelos orientados

⁷ <https://cidoc.mini.icom.museum/working-groups/lido/what-is-lido/>

⁸ <https://www.loc.gov/ead/>

⁹ <https://www.loc.gov/marc/umb/>

¹⁰ <https://www.loc.gov/standards/vracore/>

¹¹ <https://dublincore.org/>

¹² <https://mpeg.chiariglione.org/standards/mpeg-7>

por ontologias, e conhecidos na literatura por *Simple Knowledge Organization Systems (SKOS)*¹³ (GONZÁLEZ, 2011). Um exemplo seria o uso de linguagens documentárias representadas pelos tesouros *Art & Architecture Thesaurus - AAT*, *Union List of Artist Names - ULAN* e *Thesaurus of Geographic Names - TGN* do Getty, pertinentes como padrões para tratamento dos dados pertencentes a coleções de acervos culturais visando normalização, qualidade e intercâmbio de suas descrições. Tipicamente descrevem pessoas, localizações, eventos e conceitos específicos relacionados à cultura, e são mantidos pelo *The Getty Research Institute*¹⁴ através de contribuições de comunidades interessadas em suas temáticas. Apresentam terminologias estruturadas em LOD, XML e tabelas relacionais endereçadas a arte, a arquitetura, e artes decorativas, além de objetos culturais, itens de arquivos e materiais bibliográficos, estando em conformidade com normas internacionais que podem melhorar a confiabilidade e o acesso a bases de dados de cultura disponíveis na rede.

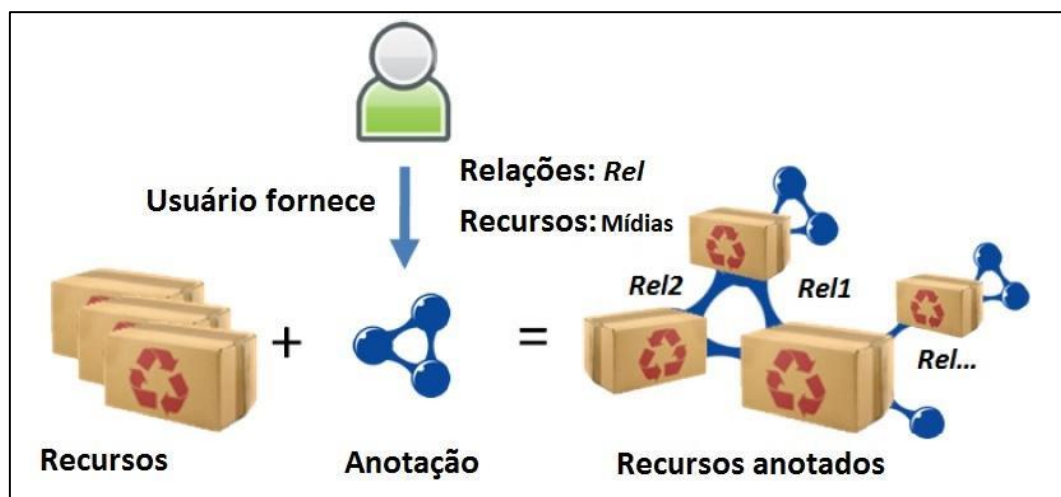
Ontologias também são usadas como suporte a anotações de forma semelhante aos vocabulários controlados, no entanto numa perspectiva de tratamento semântico, identificando formalmente conceitos e relações entre conceitos em documentos, que vão além de anotações textuais sobre o conteúdo, permitindo um usuário (humano ou não) descrever e interligar recursos existentes por meio de qualificadores como conceitos, instâncias, propriedades e restrições mantidas entre tais recursos, viabilizando o processamento pela máquina, visto que a anotação semântica deve ser explícita, formal e livre de ambiguidade, respectivamente, acessível, entendida e identificável publicamente (UREN et al., 2006; SHADBOLT; BERNERS-LEE; HALL, 2006; LEMOS; SOUZA, 2020). A Figura 2 exhibe recursos (mídias digitais) sendo anotados por elementos de ontologias, promovendo, desse modo, interligações semânticas (Rel 1, Rel 2...) entre os recursos.

¹³ <https://www.w3.org/2004/02/skos/intro>

¹⁴ <https://www.getty.edu/research/tools/vocabularies/>



Figura 1 - Modelo de anotação por meio de ontologias



Fonte: adaptado de Andrews, Zaihrayeu e Pane (2012)

Nesse caso, usuários fornecem elementos de anotação (idealmente, de forma simples e transparente) e promovem ligações dessas anotações a fontes de conhecimento subjacentes a ontologias disponíveis da Web de dados (MACHADO; SOUZA; SIMÕES, 2019). Por exemplo, para evitar ambiguidade sobre o termo “Paris”, o qual contextualiza a origem de uma fotografia digitalizada (mídia do tipo imagem) pertencente a um dado acervo, um catalogador poderia fazer uso de entidades semânticas estruturadas no Wikidata (cada uma dotada de seu *Internationalised Resource Identifier - IRI*) e, como tal, a anotação semântica poderia relacionar esse termo a um elemento da ontologia wiki que o identificasse na categoria “Cidade” bem como associá-lo à instância “França” pertencente à categoria “País”. Desse modo, “Paris” não poderia ser referenciada de outra forma a não ser como uma cidade francesa. Essa fotografia oriunda de um acervo cultural poderia então ilustrar um artigo publicado numa página do Wikipédia, manter a sua mídia atualizada no Commons juntamente com suas descrições (a nível de suporte, conteúdo e proveniência) e declarações formais mantidas coletivamente pela base de conhecimento do Wikidata, agregando o referido acervo ao ecossistema wiki por meio de ontologias.

O cenário evidencia um panorama de uso e reúso de SOCs (LEMOS; SOUZA, 2020) na organização de uma quantidade significativa de dados que tem sido gerada, conectada e distribuída na web, especialmente dados multimídia no escopo LOD e das plataformas wikis, os quais necessitam de modelos de anotação eficientes para gestão, organização, representação e recuperação de conteúdos associados a acervos culturais distribuídos em redes de conhecimento.

3 METODOLOGIA

Esta pesquisa se classifica como sendo de natureza teórica e qualitativa, de cunho exploratório e descritivo em que se vislumbrou uma revisão sistemática de literatura (RSL) (OKOLI, 2015) acerca de vocabulários e ontologias para os quais se apresentam como modelos de anotação no âmbito de projetos de integração de acervos culturais envolvendo dados abertos ligados e redes de informação da Fundação Wikimedia. Para tal, como procedimento técnico de coleta e análise dos dados, utilizou-se de pesquisa bibliográfica e documental, incluindo artigos científicos, livros, anais de eventos científicos de referência, bem como fontes documentais a partir de sítio web, manuais, tutoriais, vídeos explicativos, entre outros conteúdos associados à investigação.

A principal base de dados de documentos científicos utilizada nesta pesquisa foi o Portal de Periódicos Capes para recuperar artigos recentes, relevantes e com fator de impacto considerável aos propósitos de uma RSL. A fim de aperfeiçoar os resultados obtidos, foram consultadas diretamente quatro bases de dados de pesquisa acadêmica *online*, a saber: i) Association Computing Machinery, conhecida como portal ACM Digital Library, ii) Emerald; iii) Science Direct; e iv) Springer Link. A primeira dissemina literatura de diversas áreas da CC, destacando representação do conhecimento e engenharia de ontologias, áreas nas quais há várias publicações sobre ontologias e suas aplicações. A segunda foi escolhida por oferecer uma coleção de periódicos e livros de relevância para a CI, além de abranger áreas do conhecimento correlatas. A terceira e a quarta dão acesso a revisões sobre ferramentas e aplicações emergentes envolvendo anotações de documentos multimídia com uso de vocabulários semânticos, especialmente ontologias.

Para a recuperação dos documentos nessas bases de dados científicas foi empregada a técnica de busca por palavras-chave que refletem o universo do assunto. As pesquisas consideraram em sua totalidade apenas artigos completos em idiomas inglês, espanhol ou português, publicados em conferências e periódicos que incluíam os termos "linked open data" AND "semantic" AND "wikidata" AND "ontolog*" AND "cultur*" em qualquer parte de seu conteúdo. Um total preliminar de 97 artigos recuperados (Tabela 1) nas bases de dados de pesquisa acadêmica *online*, que, após a retirada das duplicatas, foi reduzido para 89.



Tabela 1 – Quantidade de resultados por serviço de busca

FONTE	QUANTIDADE DE RESULTADOS
ACM DL	10
Emerald	11
Portal CAPES	56
Science direct	11
Springer link	9

Fonte: Elaborado pelos autores (2021)

Todos os itens foram anotados e verificados manualmente para determinar sua relevância; uma série de outros critérios foram especificados para selecionar os estudos apropriados para inclusão na revisão. Para serem incluídos, os artigos precisariam discorrer sobre a aplicação prática de ontologia em acervo de instituições culturais, removendo revisões de literatura e capítulos de livros; fazer o uso de ontologias; realizar o enriquecimento ou integração destes acervos em ambiente LOD ou envolvendo o ecossistema wiki; além de serem publicados em periódicos ou anais de conferências e estarem disponíveis para acesso. Como resultado dessa seleção, um total de 16 artigos que atenderam a estes critérios foram selecionados para a revisão (Tabela 2).

Tabela 2 – Passos de inclusão dos artigos selecionados

FASE	CRITÉRIO	ARTIGOS RESTANTES
Resultados da busca	Artigo completos em inglês, espanhol ou português sem duplicatas	89
Exclusão de estudos que não atendem critérios	Estar acessível	73
	Não são revisão de literatura ou pesquisa de opinião ou capítulos de livros	
	Tratam de instituições de acervos culturais Fazer o uso de ontologias	23
	Realizar enriquecimento ou integração a partir do Wikidata ou com outros <i>datasets</i> em ambiente LOD.	16

Fonte: Elaborado pelos autores (2021)

De modo a organizar a análise e os resultados acerca das ontologias elencadas no âmbito da pesquisa, foram determinadas categorias de análise à luz dos estudos desenvolvidos em Lemos e Souza (2020) em que se analisa comparativamente ontologias no domínio multimídia cuja fundamentação se dá a partir de três tipologias de metadados que caracterizam recursos de



mídia, a saber: independentes de conteúdo (metadados para gerenciamento e administração de recursos de informação), dependentes de conteúdo (metadados visuais e para áudio, ambos considerados de nível primitivo e, geralmente, seus conteúdos são extraídos automaticamente por algoritmos computacionais); e descritivos de conteúdo (metadados que associam entidades da mídia com entidades do mundo real).

As demais categorias analíticas são enumeradas em sequência: propósito de desenvolvimento, recursos de conhecimento utilizados, linguagem de implementação, equipe de desenvolvimento, projetos que fazem uso, site oficial e atualização mais recente. Acredita-se que a partir de categorias analíticas bem fundamentadas em princípios metodológicos da Engenharia de Ontologias, têm-se as condições necessárias para decisões de seleção e reuso direcionadas à modelagem de conhecimento acerca de acervos multimídia do patrimônio cultural.

4 ONTOLOGIAS PARA REPRESENTAÇÃO DO CONHECIMENTO MULTIMÍDIA DO PATRIMÔNIO CULTURAL

Um objeto multimídia contempla um documento composto com o qual se faz referência a vários tipos de objetos, tais como vídeo, texto, som, imagem, modelo 3D, entre outros, e que ainda pode ser dividido em partes que resultam em tipos de mídias específicos (LEMOS; SOUZA, 2020). Assim sendo, do ponto de vista da representação de um objeto digital, deve-se levar em consideração características que vão além dos tradicionais dados clássicos ou convencionais compostos de símbolos alfanuméricos, representando nomes, códigos, medidas, quantidades e valores (SILBERSCHATZ; SUNDARSHAN; KORTH, 2016), isto é, características multimídia que são peculiar (ADJEROH; NWOSU, 1997; BARRETO, 2007), tais como relações espaciais entre elementos de interesse dentro do conteúdo (para imagem, modelo 3D e vídeo) e relações temporais na ocorrência de eventos dentro de um período de tempo (para vídeo, áudio e sequências de animação).

Considerando, portanto, a complexidade em se descrever características de objetos multimídia, entende-se a necessidade, a priori, de organizar e representar aspectos associados ao suporte ou à mídia (exs.: formato, localização, direito e licença de uso), a conteúdo documental do patrimônio cultural (exs.: cor, textura, forma, movimento, sinais de áudio, evento, objeto, agente, lugar, tempo), bem como a aspectos ligados a proveniência (exs.: origem e métodos de produção do objeto digital), a fim de se obter uma documentação condizente à realidade do domínio, pensando em acesso, recuperação e reuso. Logo, os modelos ontológicos elencados e elucidados



a seguir fazem parte desse conjunto de vocabulários que estão sendo usados e reusados em variados contextos e necessidades no âmbito de projetos associados ao patrimônio cultural digital realizado em vários tipos de mídias. Nesse sentido, alguns projetos pontuais envolvendo instituições culturais que estão praticando reuso desses padrões de documentação são brevemente apresentados a seguir.

4.1 Modelo conceitual de referência do patrimônio cultural e suas extensões

O modelo conceitual de referência CIDOC-CRM (BOEUF et al., 2018) foi criado no ano de 1994 pelo Comitê Internacional de Documentação (CIDOC) pertencente ao Conselho Internacional de Museus (ICOM) e aceito em setembro de 2006 como padrão (ISO 21127). Representa uma ontologia formal em OWL DL, incluindo uma estrutura taxonômica central ou de núcleo (*core*, do inglês) para o domínio do patrimônio cultural com características de alto nível na medida em que lida com propriedades (em torno de 200), incluindo relações temporais, de causalidade e mereológicas; além de classes/entidades gerais (em torno de 100), incluindo espaço, tempo, matéria, pessoa, objeto, documento, evento, ação, entre outras, o que facilita a integração e o intercâmbio de informações heterogêneas do patrimônio cultural. Duas extensões do modelo CIDOC-CRM para documentação de objetos digitais são elucidadas a seguir como propostas mais simplificadas para documentação do patrimônio cultural, a saber: o CRMdig (DOERR et al., 2016) e o Linked Art Data Model (FINK, 2018).

O CRMdig é um modelo conceitual ontológico para organização semântica de metadados voltados à anotação de cenários de proveniência, especialmente envolvendo modelos 3D. O modelo possui a estrutura codificada em RDF *Schema* e é uma extensão da ontologia do CIDOC-CRM, sendo capaz de capturar os requisitos de modelagem e de consultas relacionados à proveniência de objetos digitais (incluindo etapas e métodos de produção) para um domínio, em especial o de patrimônio cultural. O CRMdig consegue descrever as circunstâncias físicas envolvendo o objeto em estudo, resultando em dados digitais. Lida com informações semânticas para monitorar, gerenciar e documentar as origens e as derivações de produtos de maneira flexível. O modelo é considerado maduro em termos de cobertura, extensibilidade, granularidade e poder expressivo em lidar com dados interligados.

O Linked Art é um modelo de dados simplificado do CIDOC-CRM criado pelo American Art Collaborative com o propósito de disponibilizar e compartilhar coleções de instituições americanas no ambiente LOD. Sendo assim, também se configura como um modelo conceitual ontológico, representado em RDF *Schema* no formato *linked data* JSON-LD, destinado a



descrever obras de arte ou objetos de museus de um modo operacionalmente mais prático no que diz respeito ao manuseio de classes e relações em comparação ao uso da complexa estrutura da ontologia CRM. Nesse sentido, o modelo torna-se mais atrativo no quesito de estimular a inserção de instituições culturais no contexto de LOD e da Web semântica, favorecendo assim a estratégia de interoperabilidade entre diferentes tipos de acervos.

Considerando que o CRMdig é uma extensão do CIDOC CRM que, por sua vez, também propõe um modelo de dados simplificado no contexto de dados abertos interligados, a ontologia Linked Art, torna-se possível estruturar e interligar variados tipos de metadados de natureza multimídia em classes fundamentais de suas estruturas, tais como *Information Object*, *Document*, *Person*, *Object*, *Activity*, *Place* e *Time*, tendo em vista a conceituação comum concebida para os três modelos conceituais ontológicos. A organização de metadados em categorias fundamentais torna-se útil no esclarecimento do significado pretendido por meio de um conjunto de distinções semânticas, evitando ambiguidade e melhorando, principalmente, a qualidade na representação da informação nos acervos de cultura.

Em Metilli, Bartalesi e Meghini (2019) é apresentada uma aplicação semi-automática para biblioteca digital em que se constrói uma ontologia de narrativas objetivando criar uma rede temporal de eventos indexados (uma *timeline*) para um dado assunto de interesse do usuário, incluindo pesquisadores, alunos e professores. A modelagem ontológica consistiu no reuso do modelo de referência do CIDOC-CRM para a organização de entidades relacionadas ao assunto da narrativa trabalhada (ex.: a vida de Dante Alighieri), as quais são importadas da base de conhecimento do Wikidata com suas possíveis mídias correspondentes do Commons, sendo classificadas como instâncias na classe de eventos do CRM, que, por sua vez, possibilita agregar e classificar pessoas, objetos (físicos e abstratos), localidades e intervalos de tempo, estruturando semanticamente a narrativa construída.

Na Arqueologia, Catalano et al. (2020) apresentam um modelo de dados semântico para documentar fragmentos de artefatos culturais 3D com base no CIDOC CRM e no CRMdig buscando documentar principalmente a proveniência dos artefatos digitais, incluindo a aquisição digital e o processamento dos objetos físicos, representando características e medições precisas desses artefatos em seus aspectos geométricos que podem vir a ser úteis à reconstrução de artefatos a partir de fragmentos anotados semanticamente, à computação gráfica em geral e a ferramentas de análise de formas. Enfatizam que no domínio curatorial, coleções digitais anotadas por ontologias como o CRMdig podem se beneficiar de uma



documentação consistente e certificada para fins de exposição ao público. No caso de museus e exposições virtuais, para o público em geral, os artefatos digitais podem ser exibidos juntos com um rico conjunto de informações e descobertas arqueológicas.

O museu de médio porte dedicado à artista Georgia O'Keeffe (NEELY; LUTHER; WEINARD, 2019) adotou o modelo Linked Art para a modelagem mais simplificada de dados de seus sistemas de informação visando a abertura de suas coleções de arte, arquivo e biblioteca a fim de serem compartilhadas e agregadas a outros museus ao redor do mundo a partir dos preceitos LOD. Para tal, a partir do uso do Linked Art, aspectos de usabilidade foram levados em consideração na produção e no acesso ao conhecimento sobre a artista.

4.2 Modelo conceitual para dados bibliográficos e suas extensões

O modelo conceitual para informação bibliográfica FRBR foi desenvolvido entre 1992 e 1997 e publicado em 1998 pelo grupo de estudo do International Federation of Library Associations and Institutions (IFLA). O modelo é fundamentado na abordagem entidade e relacionamento (FRBRer) de Peter Chen (CHEN, 1976) apresentando conceitos e definições de entidades, relacionamentos e atributos, inovando o universo bibliográfico com ênfase no usuário e suas ações de encontrar, identificar, selecionar e obter o objeto bibliográfico (IFLA, 2009).

O modelo é composto por 10 entidades, divididas em três grupos, que podem ser de complexa compreensão por possuírem entidades consideradas concretas e abstratas, a saber: i) Grupo 1: entidades que representam o produto de trabalho intelectual e artístico - Obra (*Work*), Expressão (*Expression*), Manifestação (*Manifestation*) e Item (*Item*); ii) Grupo 2: entidades responsáveis pelo conteúdo intelectual, pela guarda e disseminação das entidades do Grupo 1 - Pessoa (*Person*) e Entidade Coletiva (*Corporate Body*); e iii) Grupo 3: entidades que servem como descritores de assunto da entidade Obra do Grupo 1 - Conceito (*Concept*), Objeto (*Object*), Evento (*Event*) e Lugar (*Place*).

Nesse sentido, o FRBR promove uma percepção mais abrangente sobre o item bibliográfico, permitindo uma distinção entre o conteúdo (Obra e Expressão) e a forma física do item (Manifestação e Item), e ainda possibilita agrupamentos por semelhanças e relacionamentos (com restrições de cardinalidade) entre as entidades redundando em benefícios como: i) eficácia na recuperação da informação; ii) visualização facilitada de diferentes expressões de uma obra; iii) facilidade em encontrar todas as manifestações; e iv) clareza no entendimento dos relacionamentos entre as entidades dos Grupos 1, 2 e 3, promovendo uma comunicação sob



o mais alto nível de abstração a respeito do universo bibliográfico, viabilizando, em função disso, a estratégia de busca e a recuperação da informação por parte dos usuários.

Atualmente, o modelo FRBR encontra-se harmonizado com a ontologia de alto nível do CIDOC CRM, um esforço de ambas as comunidades para oferecer suporte à integração e interoperabilidade de dados para as bibliotecas e os museus, o que resultou na versão do modelo denominada FRBRoo¹⁵ (de orientado a objeto). A segunda versão do modelo contempla além do FRBR, os modelos *Functional Requirements for Authority Data* (FRAD) para organização de entidades específicas de dados de autoridade como Nome, Identificador, Ponto de Acesso Controlado, Regras e Agência, e *Functional Requirements for Subject Authority Data* (FRSAD) para organização de dados de autoridade de assunto. O modelo FRBRoo também se apresenta corrigido em relação a inconsistências apresentadas no antigo modelo (ex. a entidade Manifestação passou a ser representada em duas classes distintas: uma com visão conceitual e outra relacionada à visão física), além da inserção da noção de Evento no mesmo.

Candela et al. (2019) relata uma experiência com a Biblioteca Virtual Miguel de Cervantes em que os dados bibliográficos de textos em espanhol, ou relacionado à cultura hispânica, são enriquecidos a partir de metadados estruturados do Wikidata e do GeoNames, mostrando como o enriquecimento semântico em padrões de catalogação tradicionais, como MARC, pode ajudar na recuperação da informação de itens bibliográficos na rede. O modelo ontológico da aplicação de biblioteca digital foi concebido a partir das entidades bibliográficas do modelo FRBR fundamentadas na norma de catalogação emergente *Resource Description and Access* (RDA) ou Descrição e Acesso de Recursos.

Nessa mesma linha de integração com o Wikidata, a ontologia Ontophoto (ROBLEDANO-ARILLO; NAVARRO-BONILLA; CERDÁ-DÍAZ, 2019) importa dessa base de conhecimento, entidades endereças à descrição da imagem em seus aspectos de: características físicas, de autoria e organizacionais das fotografias; atributos de conteúdo iconográfico; atributos de conteúdo temático; características legais; e conhecimento no domínio temático retratado nas imagens (ontologia de domínio). Dentre os vários esquemas externos reusados, incluindo o EDM, a ontologia possui em sua estrutura taxonômica, elementos do modelo FRBR para a organização das entidades bibliográficas (instâncias) em grupos documentais coerentes e,

¹⁵ <http://www.cidoc-crm.org/frbroo/>

desse modo, difundir uma documentação consistente dos acervos históricos fotográficos na rede.

4.3 Modelo conceitual do Europeana Collection

O EDM (*Europeana Data Model*) é a especificação formal das classes e propriedades que podem ser utilizadas no portal Europeana Collections. O modelo teve sua primeira versão lançada em fevereiro de 2009 pela *Europeana Foundation*, contando com a colaboração de especialistas técnicos das áreas de acervos de bibliotecas, museus, arquivos e do audiovisual na adoção dos preceitos da Web Semântica, cujo formato adotado é o *RDF Schema*.

Como base para a construção do EDM foram utilizados o OAI ORE (*Open Archives Initiative Object Reuse and Exchange*) na organização dos metadados do objeto e de suas representações digitais, o Dublin Core para os metadados descritivos, o SKOS para a representação conceitual dos contextos relacionados ao conteúdo do objeto e o CIDOC CRM para os eventos e relacionamentos entre os objetos.

O EDM apresenta sete classes, sendo que três representam o artefato cultural e quatro contextuais que podem ser associadas aos objetos, a saber: i) *edm:ProvidedCHO* (*Provided Cultural Heritage Object*), que descreve o objeto físico; ii) *edm:WebResource*, que descreve a representação digital; iii) *ore:Aggregation*, que agrupa as entidades de relacionamento. Por sua vez, as entidades de relacionamento basilares entre as três classes apresentadas anteriormente são a *edm:aggregatedCHO* e a *edm:hasView*. Já as classes contextuais são: i) *edm:Agent*, que vincula ao agente, ii) *edm:Place*, que vincula ao local, iii) *edm:TimeSpan*, que vincula a um período do tempo e iv) *skos:Concept*, que vincula ao conceito (CARRASCO; VIDOTTI, 2018).

Alexiev e Georgiev (2020) relatam sobre a exportação da Enciclopédia Virtual de Ícones Búlgaros para o Wikidata de forma a enriquecer com informações extras as entidades de sua coleção. Os autores relatam que no processo de mapeamento dos itens do acervo com a ontologia EDM foi feito o enriquecimento dos recursos com a inclusão de *links* a fontes externas referenciadas no Wikidata com a inclusão de descritores bilíngues ao acervo. Após o tratamento de curadoria nos dados, a base foi então enviada à biblioteca digital Europeana para a difusão do acervo na rede de agregação.



Beck et al. (2017) discorre sobre a migração do acervo de artes cênicas disponível em um portal de busca de forma a serem ligados a *datasets* da nuvem LOD, permitindo disponibilidade e reúso. Para isso, o padrão de metadados MARC21 utilizado no acervo foi substituído pela ontologia EDM, pois, conforme os autores, o padrão MARC21 não era apropriado para mapear os metadados de museus e arquivos nos preceitos LOD. Assim, com a adoção da ontologia EDM, foi possível integrar os dados do acervo a um *dataset* LOD da Alemanha chamado *Entity Facts* que fornece conjuntos de fatos sobre entidades do tipo pessoas e corporações catalogadas na literatura em bibliotecas de países falantes de Alemão. Posteriormente, esses dados foram ligados a entidades do Wikidata visando enriquecimento e construção coletiva de conhecimento para o acervo.

O Quadro 1 a seguir sintetiza o resultado das análises realizadas nos modelos conceituais ontológicos com o propósito de orientar em decisões de uso e reúso envolvendo aspectos de suas estruturas a partir de requisitos funcionais e não funcionais (LEMOS, 2020) pré-estabelecidos.



Quadro 1 – Quadro sinóptico dos modelos ontológicos elencados

CATEGORIAS DE ANÁLISE		CIDOC CRM	CRMDIG	LINKED ART	EDM	FRBR E EXTENSÕES
Propósito de desenvolvimento		Facilitar a integração, mediação e intercâmbio de informações heterogêneas do patrimônio cultural na rede.	Organizar semanticamente metadados voltados à documentação de cenários de proveniência.	Servir como uma adaptação mais simples e mais fácil de implementação baseada no modelo CIDOC CRM, facilitando a entrada de instituições no mundo dos dados abertos ligados.	Enriquecer dados e metadados culturais por meio de estruturas de alto nível, servindo como agregador de acervos culturais em LOD.	Capturar e representar a semântica subjacente a informações bibliográficas, facilitando a integração, a mediação e o intercâmbio de dados bibliográficos na rede.
Natureza dos metadados multimídia	Independentes de conteúdo	Criação, produção, classificação, informação técnica e direito de uso do recurso original e digital.	Criação, produção e classificação do recurso original e digital.	Criação, produção, classificação, informação técnica e direito de uso do recurso original e digital.	Criação, produção, classificação, informação e uso do recurso.	Grupo 1: entidades que representam o produto de trabalho intelectual e artístico associado à criação e à produção da obra.
	Dependentes de conteúdo	Não identificado.	Elementos de descrição quantitativa relacionada a medição digital.	Não identificado.	Não identificado.	Não identificado.
	Descritivos de conteúdo	Modelo centrado em entidades abstratas para representação de um objeto de informação e sua realização em mídias.	Anotações para o evento da aquisição de dados e do objeto digital em si no todo ou em suas partes.	Modelo centrado em entidades abstratas para representação de um objeto cultural.	Classes contextuais para descrever conteúdo sobre o recurso (edm:Agent, edm:Place, edm:TimeSpan, skos:Concept).	O Grupo 3 define quatro entidades denominadas Conceito, Objeto, Evento e Lugar.
Recursos de conhecimento utilizados		Não identificado.	CIDOC-CRM.	CIDOC-CRM, Getty Vocabularies.	Dublin Core, FOAF, MARC, BIBO, EAD, LIDO, SKOS, DCAT (Data Catalog Vocabulary), o CC (creative commons)	Modelo entidade e relacionamento, CIDOC-CRM.
Linguagem de implementação		OWL DL.	RDF Schema.	RDF Schema, JSON-LD.	RDF Schema, RDF/XML.	OWL DL.

Equipe de desenvolvimento	Equipe multidisciplinar sob o amparo do CIDOC do ICOM.	Comitê CIDOC.	American Art Collaborative.	Europeana Foundation.	IFLA.
Projetos que fazem uso	University of Pisa (METILLI; BARTALESI; MEGHINI, 2019); Catania's Benedictines Monastery (CANTALE et al., 2017); Zeri Photo Archive catalog (DAQUINO et al. 2017); Dunhuang cultural heritage (WANG; CHANG; TAN, 2020)	Gravitate (CATALANO et al., 2020)	American Art Collaborative (FINK, 2018); Museu Georgia O'Keeffe (NEELY; LUTHER; WEINARD, 2019)	Rijksmuseum Collection (DIJKSHOORN et al., 2018); Arquivo fotográfico Skogler (ROBLEDANO-ARILLO; NAVARRO-BONILLA; CERDÁ-DÍAZ, 2019); Specialised Information Service Performing Arts (BECK et al., 2017); Virtual Encyclopedia of Bulgarian Icons (ALEXIEV; GEORGIEV, 2020)	Arquivo fotográfico Skogler (ROBLEDANO-ARILLO; NAVARRO-BONILLA; CERDÁ-DÍAZ, 2019); Princeton University Library (HEBERLEIN, 2019); Biblioteca Virtual Miguel de Cervantes (CANDELA et al., 2019); Baden-Württemberg State Archives (VSESVIATSKA et al., 2021); Dunhuang cultural heritage (WANG; CHANG; TAN, 2020)
Site oficial	cidoc-crm.org	cidoc-crm.org/crmdig/	linked.art	pro.europeana.eu	iifla.org
Atualização mais recente	maio, 2021	abril, 2016	junho, 2021	outubro, 2017	novembro, 2015

Fonte: Elaborado pelos autores (2021)



5 DISCUSSÕES

Os cinco modelos ontológicos apresentados, o CIDOC CRM, suas extensões CRMdig e Linked Art, o FRBRoo e o EDM mostram-se conceitualmente, em tese, adequados à modelagem de documentos realizados em vários tipos de suportes e mídias no âmbito da cultura por apresentarem em suas estruturas taxonômicas, classes e relacionamentos de alto nível baseados em aspectos cognitivos, filosóficos e linguísticos (GUIZZARDI, 2005) para a descrição formal de obras, objetos, eventos, tempo, espaço, organizações, entre outros elementos centrais para descrição semântica de documentos multimídia em seus aspectos de proveniência, de mídia e de conteúdo do patrimônio cultural.

A semântica formal é capturada pelos formatos padrão W3C serializados em RDF/XML, RDF Schema e OWL DL, sendo essa última mapeada em lógica descritiva com suporte a inferências lógicas, dando, portanto, condições necessárias ao tratamento sintático e semântico dos dados envolvidos em seus esquemas (ALLEMANG; HENDLER; GANDON, 2020). Nesse quesito, as ontologias do CIDOC CRM e do FRBRoo se sobressaem em relação aos demais modelos ora analisados.

As características funcionais estabelecidas na pesquisa para metadados multimídia (independentes, dependentes e descritivos de conteúdo) destinados à representação de documentos dessa natureza se mostram bem representativas nos modelos conceituais analisados, porém nota-se uma carência de metadados primitivos visuais (cor, textura, forma, movimento e localização de regiões) e de áudio (sinais e arquivos) associados à categoria dependentes de conteúdo, com exceção da ontologia do CRMdig que apresenta categorias formais em sua estrutura ontológica para aspectos de medição digital com propriedades quantitativas relacionadas a entidades físicas envolvidas em processos de digitalização, representando características e valores livres de ambiguidades para anotações dessa natureza, abrindo caminhos, por exemplo, para uma documentação complexa, estruturada e rica como é o caso de patrimônio cultural tangível 3D (ex.: artefatos arqueológicos, esculturas, edifícios). A anotação semântica de artefatos 3D com medições precisas cria contrapartes confiáveis, tornando-se possível pesquisar e recuperar dados heterogêneos de maneira integrada e, conseqüentemente, comparar artefatos e suas características por meio de algoritmos computacionais (CATALANO et al., 2020).



Logo, considera-se viável para o preenchimento dessa lacuna, a seleção e o reuso de ontologias multimídia bem colocadas em um *ranking* obtido a partir de um estudo minucioso realizado em Lemos e Souza (2020) que traz propostas ontológicas fundamentadas em padrões de projeto multimídia (anotação, decomposição e coleção) juntamente com classes de tipos de metadados que preveem uma clara separação de interesses em relação à mídia, a saber: semântica de conteúdo, conhecimento relacionado à gestão de recursos informacionais, segmentação de mídia e propriedades quantitativas associadas a conteúdo extraído da própria mídia.

As classes e propriedades que refletem a criação, a produção, a classificação, as informações de natureza técnica da mídia e o uso do recurso original e digital se propõem a organizar ontologicamente metadados relacionados à gestão de recursos informacionais (independentes de conteúdo). Nesse caso, todos os modelos ontológicos analisados contemplam tais características por meio das quais podem ser aplicadas à realização da mídia e de seu conteúdo, promovendo, assim, a organização de descrições de alto nível a variados domínios que necessitam desta distinção semântica para representação e recuperação de recursos de informação, como é o caso, por exemplo, do universo bibliográfico, cujo modelo conceitual padrão é o FRBR para uma percepção mais abrangente sobre o item bibliográfico em relação à distinção semântica do objeto físico ou digital e suas possíveis realizações.

Já as anotações sobre o conteúdo da mídia descrevem objetos, agentes, eventos e noções do mundo real (metadados descritivos de conteúdo) que podem ser abstraídos do conteúdo e ligados a sistemas de organização do conhecimento (HJØRLAND, 2003) como ontologias, vocabulários controlados e SKOS de domínios específicos disponíveis para uso e reuso em rede de dados abertos ligados.

Nos aspectos que tangem à proveniência de dados, a estrutura taxonômica do modelo CRMdig dispõe de classes representativas de objetos e eventos (além de outras entidades interativas, como pessoa, lugar, tempo, etc) envolvidos na produção de artefatos digitais, incluindo dispositivos, algoritmos computacionais e pessoas que participam do fluxo de trabalho. Em função do alto nível das categorias representadas na estrutura taxonômica do modelo, os eventos relacionados aos metadados de proveniência podem ser estruturados, respondendo a quatro perguntas principais, a saber: i) as pessoas ou organizações que participaram do evento; ii) o local do evento; iii) a data e a hora em que o evento ocorreu; iv) as entidades envolvidas no evento; e v) o tipo de processo e técnicas aplicadas. Essas informações são fundamentais ao tratamento, por exemplo, de réplicas digitais de objetos culturais 3D (LEMOS; MEDEIROS e SÁ;

SOUZA, 2019) no sentido de garantir a transparência da relação entre a réplica digital e o objeto físico real, além de possibilitar repetibilidade e verificabilidade sobre o experimento.

A confiabilidade no quesito reúso pode ser considerada uma característica presente em todas as ontologias analisadas, especialmente o CRM com seu padrão ISO, e pelo fato de possuírem uma equipe de desenvolvimento com boa reputação (CIDOC do ICOM, American Art Collaborative, Europeana Foundation e IFLA) e assistidas por importantes entidades no cenário mundial como é o caso do W3C. Adicionalmente, grande parte dos vocabulários analisados é fundamentada em recursos de conhecimento (padrões de metadados, vocabulários controlados, modelos conceituais e ontologias) consolidados na comunidade da qual fazem parte, reforçando a confiabilidade no uso e reúso (LEMOS; SOUZA, 2020) frente a aspectos de tratamento semântico nas declarações. Isso se confirma nos projetos recentes levantados na revisão de literatura em variados contextos e necessidades no uso desses padrões de documentação endereçados à organização e representação de acervos museais, bibliográficos e arquivísticos.

Por fim, os modelos ontológicos analisados estão até o presente momento desta pesquisa com suas fontes de dados acessíveis e atualizadas, conforme indica o Quadro 1.

6 CONSIDERAÇÕES FINAIS

Os vocabulários selecionados, analisados e discutidos na presente pesquisa, mesmo com uma análise ainda não aprofundada para fins de alinhamentos práticos previstos em uma fase metodológica relacionada à modelagem conceitual à luz de requisitos funcionais e não funcionais pré-estabelecidos (LEMOS, 2020), mostram-se como bons modelos ontológicos candidatos a reúso, sobretudo em projetos que visem a integração de acervos culturais em ambiente LOD e envolvendo o ecossistema wiki.

Isso se deve ao fato de esses modelos possuírem em suas estruturas centrais, classes e relacionamentos considerados de alto nível ontológico e que são capazes, portanto, de descrever formalmente objetos digitais multimídia pertencentes a metacategorias como eventos, objetos, tempo, espaço, apenas para citar algumas, para organizar semanticamente conteúdos advindos de vocabulários e ontologias para variados contextos específicos do domínio do patrimônio cultural. Nesse sentido, a semântica formal oriunda dos padrões da Web semântica com os quais suas estruturas são representadas, além de prever meios de transmissão em alguma sintaxe acordada pela comunidade, assegura que o significado

intencionado da semântica ligada aos dados das bases das instituições culturais possa ser compartilhado entre diferentes aplicações no escopo dos dados abertos ligados e do ecossistema wiki.

O objetivo do artigo cumpre-se, portanto, na medida em que apresenta uma análise sistemática, mesmo que simplificada, do CIDOC CRM, do CRMdig, do Linked Art, do EMD e do FRBRoo, respondendo às questões previstas na pesquisa em saber como esses padrões de documentação estão sendo usados como formas de representação do conhecimento multimídia no contexto do patrimônio cultural e como organizam em suas estruturas, metadados de natureza multimídia (independentes, dependentes e descritivos de conteúdo) e de proveniência (origem e métodos de produção) para descrição de objetos culturais em variados contextos e necessidades.

Finalmente, considerando que instituições no âmbito GLAM têm se tornado importantes provedores de dados multimídia nas redes de informação do Wikidata, do Wikimedia Commons e do Wikipédia, conforme se apresentou em alguns casos relatados nesta pesquisa, especialmente em continente europeu, as anotações semânticas providas de suas ontologias se mostram vantajosas ao permitir relações explícitas e significativas entre os elementos estruturados das ontologias e os metadados e mídias associadas das bases de dados ora envolvidas nessa rede de conhecimento colaborativa. Isso possibilita uma variedade de técnicas de recuperação da informação baseadas em esquema de conhecimento expresso em uma ontologia, como raciocínio automatizado, coocorrência de anotação ou entidades no mesmo recurso ou contexto, potencializando, inclusive, retroalimentação de dados oriundos da catalogação original dos objetos culturais a partir de sugestões, correções e revisões realizadas por usuários humanos bem como robôs.

Como mencionado anteriormente, é possível notar uma carência de projetos e iniciativas no Brasil no que se refere à descrição de coleções e do conhecimento subjacente das instituições nacionais de patrimônio cultural no ecossistema wiki. Entretanto, esse panorama tem a tendência de mudar progressivamente com algumas iniciativas atuais, tais como a presença de acervos de museus ligados ao Ibram no Wikimedia Commons e de suas páginas na rede do Wikipédia (CARMO; MARTINS, 2020), além de vários registros de imagens e artigos também publicados no Commons e na Wikipédia pelo Arquivo Nacional e pelos museus Paulista e de Anatomia Veterinária, com algumas iniciativas de publicação no Wikidata (PESCHANSKI, 2021).

Nesse caminho, tem-se como encaminhamento de pesquisa futura a construção de um modelo conceitual ontológico que sirva de base para a agregação de acervos culturais brasileiros nas



redes de informação das plataformas wikis no ensejo de ampliar a difusão das coleções patrimoniais, de tornar explícito o conhecimento gerado em seus acervos, e, logo, valorizar e democratizar a cultura brasileira na sociedade em rede. Para tal, mirando a estratégia de reúso dos modelos ontológicos ora pesquisados e a construção de ontologias próprias para representar, descrever e integrar o conteúdo de acervos culturais ao ecossistema wiki, pretende-se utilizar uma metodologia própria de construção de ontologias - a OntoForInfoScience (MENDONÇA, 2017) - e um editor web gráfico - o Onto4ALLEditor (MENDONÇA et al., 2020) -, que permitam uma melhor experiência do usuário (UX - User eXperience) no desenvolvimento de suas próprias ontologias.

REFERÊNCIAS

- ALEXIEV, Vladimir; GEORGIEV, Nikola. Bulgarian Icons in Wikidata and EDM. In Digital Presentation and Preservation of Cultural and Scientific Heritage (DIPP 2020). **Institute of Mathematics and Informatics (IMI BAS)**, v. 10, p. 45–63, 2020.
- ABBAS, June. **Structures for Organizing Knowledge: Exploring Taxonomies, Ontologies, and Other Schema**. New York: Neal-Schuman Publisher, 2010.
- ADJEROH, Donald A.; NWOSU, Kingsley C. Multimedia Database Management & Requirements and Issues. **IEEE MultiMedia**, v. 4, n. 3, p. 24–33, 1997.
- ALLEMANG, Dean; HENDLER, Jim; GANDON, Fabien. **Semantic Web for the Working Ontologist: Effective Modeling for Linked Data, RDFS, and OWL**. 3. ed. New York, NY, USA: Association for Computing Machinery, 2020.
- ALMEIDA, Mauricio Barcellos. Revisiting ontologies: A necessary clarification. **Journal of the American Society for Information Science and Technology**, v. 64, n. 8, p. 1682–1693, 2013.
- ANDREWS, Pierre; ZAIHRAYEU, Ilya; PANE, Juan. A classification of semantic annotation systems. **Semantic Web**, v. 3, n. 3, p. 223–248, 2012.
- BARRETO, Juliano Serra. Desafios e avanços na recuperação automática da informação audiovisual. **Ciência da Informação**, v. 36, n. 3, p. 17–28, 2007.
- BECK, Julia; BÜCHNER, Michael; BARTHOLMEI, Stephan; *et al.* Performing Entity Facts: The Specialised Information Service Performing Arts. **Datenbank-Spektrum**, v. 17, n. 1, p. 47–52, 2017.

BERNERS-LEE, TIM; HENDLER, JAMES; LASSILA, ORA. Scientific American: Feature Article: The Semantic Web: May 2001. **Scientific American**, v. 284, n. 5, p. 34–43, 2001.

BIZER, Christian; HEATH, Tom; BERNERS-LEE, Tim. Linked Data - The Story So Far. **International Journal on Semantic Web and Information Systems (IJSWIS)**, v. 5, n. 3, p. 1–22, 2009.

BOEUF, Patrick Le; DOERR, Martin; ORE, Christian Emil; *et al.* **Definition of the CIDOC Conceptual Reference Model**. Version 6.2.3 | CIDOC-CRM. Disponível em: <http://www.cidoc-crm.org/Version/version-6.2.3>. Acesso em: 30 jul. 2021.

CANDELA, Gustavo; ESCOBAR, Pilar; CARRASCO, Rafael C; *et al.* A linked open data framework to enhance the discoverability and impact of culture heritage. **Journal of Information Science**, v. 45, n. 6, p. 756–766, 2019.

CANTALE, Claudia; SANTAMARIA, Daniele Francesco; NICOLOSI-ASMUNDO, Marianna; *et al.* Distant Reading Through Ontologies : The Case Study of Catania’s Benedictines Monastery. **JLIS**, n. 8, 2017. Disponível em: <https://doi.org/10.4403/jlis.it-12342>. Acesso em: 27 jul. 2021.

CARMO, Danielle do; MARTINS, Dalton Lopes. Acervos Culturais Brasileiros no Repositório Wikimedia Commons: um estudo sobre o reúso e a visualização de mídias referentes a coleções de museus do Instituto Brasileiro de Museus. **Ciência da Informação**, v. 49, n. 3, 2020. Disponível em: <http://revista.ibict.br/ciinf/article/view/5463>. Acesso em: 30 jul. 2021.

CARRASCO, L. B.; VIDOTTI, S. A. B. G. Patrimônio Cultural: um panorama do modelo de dados da Europeana. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 19., 2018. Anais. Londrina: UEL, 2018. Disponível em: <http://enancib.marilia.unesp.br/index.php/XIXENANCIB/xixenancib/paper/view/1497>. Acesso em: 30 jul. 2019.

CATALANO, Chiara Eva; VASSALLO, Valentina; HERMON, Sorin; *et al.* Representing quantitative documentation of 3D cultural heritage artefacts with CIDOC CRMdig. **International Journal on Digital Libraries**, v. 21, n. 4, p. 359–373, 2020.

CHARLES, Valentine; CLAYPHAN, Robina; ISAAC, Antoine. Definition of the Europeana Data Model v5.2.8. 2017. Disponível em: https://pro.europeana.eu/files/Europeana_Professional/Share_your_data/Technical_requirements/EDM_Documentation//EDM_Definition_v5.2.8_102017.pdf. Acesso em: 30 jul. 2021.

CHEN, Peter Pin-Shan. The entity-relationship model — toward a unified view of data. **ACM Transactions on Database Systems**, v. 1, n. 1, p. 9–36, 1976.

DAQUINO, Marilena; MAMBELLI, Francesca; PERONI, Silvio; *et al.* Enhancing Semantic Expressivity in the Cultural Heritage Domain: Exposing the Zeri Photo Archive as Linked Open Data. **Journal on Computing and Cultural Heritage**, v. 10, n. 4, p. 1–21, 2017.

DIJKSHOORN, Chris; JONGMA, Lizzy; AROYO, Lora; *et al.* The Rijksmuseum collection as Linked Data. **Semantic Web**, v. 9, n. 2, p. 221–230, 2018.

DOERR, Martin; STEAD, Stephen; THEODORIDOU, Maria; *et al.* Definition of the CRMdig: An Extension of CIDOC-CRM to support provenance metadata. *In*: [s.l.: s.n.], 2016. Disponível em: https://projects.ics.forth.gr/isl/CRMext/CRMdig/docs/CRMdig_v3.2.1.pdf. Acesso em: 30 jul. 2021.

DOMINGUE, John; FENSEL, Dieter; HENDLER, James A. (Orgs.). **Handbook of Semantic Web Technologies**. Berlin Heidelberg: Springer-Verlag, 2011. Disponível em: <https://www.springer.com/gp/book/9783540929123>. Acesso em: 30 jul. 2021.

EUROPEANA TECH. Interim Analysis of EuropeanaTech AI in Relation to GLAMs Survey. 2020. Disponível em: https://pro.europeana.eu/files/Europeana_Professional/Europeana_Network/Europeana_Network_Task_Forces/Final_reports/Final_Interim_Report_AI_in_GLAMs_TF.pdf. Acesso em: 30 jul. 2021.

FINK, Eleanor E. **American Art Collaborative (AAC) Linked Open Data (LOD) Initiative, Overview and Recommendations for Good Practices**. Disponível em: <https://repository.si.edu/bitstream/handle/10088/106410/OverviewandRecommendationsAccessible.pdf>. Acesso em: 30 jul. 2021. Accepted: 2020-07-08T02:02:47Z.

GILLILAND, Anne J. **“Setting the Stage.” In Introduction to Metadata**. 3. ed. Los Angeles: Getty Publications, 2016. Disponível em: <http://www.getty.edu/publications/intrometadata/setting-the-stage/>. Acesso em: 30 jul. 2021.

GÓMEZ-PÉREZ, Asunción. Ontology Evaluation. *In*: STAAB, Steffen; STUDER, Rudi (Orgs.). **Handbook on Ontologies**. Berlin, Heidelberg: Springer, 2004, p. 251–273. (International Handbooks on Information Systems). Disponível em: https://doi.org/10.1007/978-3-540-24750-0_13. Acesso em: 30 jul. 2021.

GONZÁLEZ, José Antonio Moreiro. **Linguagens documentárias e vocabulários semânticos para a web: elementos conceituais**. DGO-Digital original. [s.l.]: SciELO – EDUFBA, 2011. Disponível em: <https://www.jstor.org/stable/10.7476/9788523218829>. Acesso em: 30 jul. 2021.

GUIZZARDI, Gian. **Ontological foundations for structural conceptual models**. 2005. Thesis (Telematics and Information Technology PhD) - Universidade de Twente, Enschede, Holanda, 2005. Disponível em: <<https://research.utwente.nl/en/publications/ontological-foundations-for-structural-conceptual-models>>. Acesso em: 30 jul. 2021.

HEBERLEIN, Regine. On the Flipside: Wikidata for Cultural Heritage Metadata through the Example of Numismatic Description. **IFLA WLIC**, p. 26, 2019.

HJØRLAND, Birger. Fundamentals of Knowledge Organization. **Knowledge Organization**, v. 30, n. 2, p. 87–111, 2003.

HJØRLAND, Birger. Semantics and Knowledge Organization. **Annual Review of Information Science and Technology**, v. 41, p. 367–405, 2007.

HJØRLAND, Birger. Theories are Knowledge Organizing Systems (KOS). **KNOWLEDGE ORGANIZATION**, v. 42, n. 2, p. 113–128, 2015.

IFLA. Functional requirements for bibliographic records. **Study Group on the Functional Requirements for Bibliographic Records**, p. 142, 2009.

KOLTAY, Tibor. Library and information science and the digital humanities: Perceived and real strengths and weaknesses. **Journal of Documentation**, v. 72, n. 4, p. 781–792, 2016.

LEMOS, Daniela Lucas da Silva; MEDEIROS e SÁ, Asla; SOUZA, Renato Rocha. PADRÕES PARA DOCUMENTAÇÃO DE RÉPLICAS DIGITAIS EM 3D. **Tendências da Pesquisa Brasileira em Ciência da Informação**, v. 12, n. 2, 2019. Disponível em: <https://ancib.org/revistas/index.php/tpbci/article/view/497>. Acesso em: 30 jul.2021.

LEMOS, Daniela Lucas da Silva. Modelagem de metadados multimídia: uma proposta ontológica baseada em reúso: an ontological proposal based on reuse. **Ciência da Informação**, v. 49, n. 3, 2020. Disponível em: <http://revista.ibict.br/ciinf/article/view/5451>. Acesso em: 30 jul. 2021.

LEMOS, Daniela Lucas da Silva; SOUZA, Renato Rocha. Knowledge Organization Systems for the Representation of Multimedia Resources on the Web: A Comparative Analysis. **KNOWLEDGE ORGANIZATION**, v. 47, n. 4, p. 300–319, 2020.

LEMOS, Daniela Lucas da Silva; SOUZA, Renato Rocha. Ontologias na representação de documentos: um panorama atual para descrição de conteúdo multimídia em rede. **Informação & Sociedade: Estudos**, v. 29, n. 4, p. 103–134, 2019.

LIU, Alan. 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.

MACHADO, Luís Miguel Oliveira; SOUZA, Renato Rocha; SIMÕES, Maria da Graça. Semantic web or web of data? a diachronic study (1999 to 2017) of the publications of tim berners-lee and the world wide web consortium. **Journal of the Association for Information Science and Technology**, v. 70, n. 7, p. 701–714, 2019.

MENDONÇA, Fabricio Martins; SOARES, António Lucas. Construindo ontologias com a metodologia ontoforinfoscience: uma abordagem detalhada das atividades do desenvolvimento ontológico. **Ciência da Informação**, v. 46, n. 1, 2017. Disponível em: <http://revista.ibict.br/ciinf/article/view/4013>. Acesso em: 30 jul. 2021.

MENDONÇA, Fabrício M; DE CASTRO, Lucas P; DE SOUZA, Jairo F; et al. Onto4AllEditor: um Editor Web Gráfico de Ontologias Direcionado a Diferentes Tipos de Desenvolvedores de Ontologias. **Proceedings of the XIII Seminar on Ontology Research in Brazil (ONTOBRAS 2020)**, p. 104–119, 2020.

METILLI, Daniele; BARTALESI, Valentina; MEGHINI, Carlo. A Wikidata-based tool for building and visualising narratives. **International Journal on Digital Libraries**, v. 20, n. 4, p. 417–432, 2019.

MORA-CANTALLOPS, Marçal; SÁNCHEZ-ALONSO, Salvador; GARCÍA-BARRIOCANAL, Elena. A systematic literature review on Wikidata. **Data Technologies and Applications**, v. 53, n. 3, p. 250–268, 2019.

NEELY, Liz; LUTHER, Anne; WEINARD, Chad. Cultural Collections as Data: Aiming for Digital Data Literacy and Tool Development. **MW19: MW 2019**, 2019. Disponível em: <https://mw19.mwconf.org/paper/cultural-collections-as-data-aiming-for-digital-data-literacy-and-tool-development/>. Acesso em: 30 jul. 2021.

NÚCLEO DE INFORMAÇÃO E COORDENAÇÃO DO PONTO BR. **TIC CULTURA 2018 - Pesquisa Sobre o Uso das Tecnologias de Informação e Comunicação nos Equipamentos Culturais Brasileiros**. São Paulo: Comitê Gestor da Internet no Brasil, 2019. Disponível em: https://cetic.br/media/docs/publicacoes/1/tic_cultura_2018_livro_eletronico.pdf. Acesso em: 30 jul. 2021.

OKOLI, Chitu. A Guide to Conducting a Standalone Systematic Literature Review. **Communications of the Association for Information Systems**, v. 37, 2015. Disponível em: <https://aisel.aisnet.org/cais/vol37/iss1/43/>. Acesso em: 29 jul. 2021.

OOMEN, Johan; AROYO, Lora. Crowdsourcing in the cultural heritage domain: opportunities and challenges. *In: Proceedings of the 5th International Conference on Communities and Technologies*. New York, NY, USA: Association for Computing Machinery, 2011, p. 138–149. (C&T '11). Disponível em: <https://doi.org/10.1145/2103354.2103373>. Acesso em: 30 jul. 2021.

ORIGLIA, Antonio; ROSSI, Silvia; MARTINO, Sergio Di; *et al.* Multiple-source Data Collection and Processing into a Graph Database Supporting Cultural Heritage Applications. *Journal on Computing and Cultural Heritage*, v. 14, n. 4, p. 55:1-55:27, 2021.

PESCHANSKI, João Alexandre. Variedades de processos de difusão digital colaborativa: descrição e análise de iniciativas GLAM-Wiki no Brasil. *Resgate: Revista Interdisciplinar de Cultura*, v. 29, p. 1–28, 2021.

RILEY, Jenn. UNDERSTANDING METADATA: WHAT IS METADATA, AND WHAT IS IT FOR? *National Information Standards Organization (NISO)*, p. 49, 2017. (NISO Primer Series).

ROBLEDANO-ARILLO, Jesús; NAVARRO-BONILLA, Diego; CERDÁ-DÍAZ, Julio. Application of Linked Open Data to the coding and dissemination of Spanish Civil War photographic archives. *Journal of Documentation*, v. 76, n. 1, p. 67–95, 2019.

RUIZ-CALLEJA, Adolfo; VEGA-GORGOJO, Guillermo; BOTE-LORENZO, Miguel L.; *et al.* Supporting contextualized learning with linked open data. *Journal of Web Semantics*, v. 70, p. 100657, 2021.

SHADBOLT, N.; BERNERS-LEE, T.; HALL, W. The Semantic Web Revisited. *IEEE Intelligent Systems*, v. 21, n. 3, p. 96–101, 2006.

SILBERSCHATZ, Abraham; SUNDARSHAN, S.; KORTH, Henry F. *Sistema de Banco de Dados*. São Paulo: Elsevier, 2016.

SILVA, Daniela Lucas da; SOUZA, Renato Rocha; ALMEIDA, Maurício Barcellos. Ontologias e vocabulários controlados: comparação de metodologias para construção. *Ciência da Informação*, v. 37, n. 3, p. 60–75, 2008.

SUÁREZ-FIGUEROA, Mari Carmen; GÓMEZ-PÉREZ, Asunción; FERNÁNDEZ-LÓPEZ, Mariano. The NeOn Methodology for Ontology Engineering. *In: SUÁREZ-FIGUEROA, Mari Carmen; GÓMEZ-PÉREZ, Asunción; MOTTA, Enrico; et al (Orgs.). Ontology Engineering in a Networked World*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, p. 9–34. Disponível em: http://link.springer.com/10.1007/978-3-642-24794-1_2. Acesso em: 30 jul. 2021.



TRIQUES, Maria Lígia; ARAKAKI, Ana Carolina Simionato. Representação de patrimônios culturais em plataformas digitais: o modelo de dados da Europeana. **Em Questão**, v. 27, n. 1, p. 185–209, 2021.

UREN, Victoria; CIMIANO, Philipp; IRIA, José; *et al.* Semantic annotation for knowledge management: Requirements and a survey of the state of the art. **Journal of Web Semantics**, v. 4, n. 1, p. 14–28, 2006.

VILLAESPESA, Elena; NAVARRETE, Trilce. Museum Collections on Wikipedia: Opening Up to Open Data Initiatives. **MW19: MW 2019**, 2019. Disponível em: <https://mw19.mwconf.org/paper/museum-collections-on-wikipedia-opening-up-to-open-data-initiatives/>. Acesso em: 30 jul. 2021.

VSESVIATSKA, Oleksandra; TIETZ, Tabea; HOPPE, Fabian; *et al.* ArDO: an ontology to describe the dynamics of multimedia archival records. *In: Proceedings of the 36th Annual ACM Symposium on Applied Computing*. Virtual Event Republic of Korea: ACM, 2021, p. 1855–1863. Disponível em: <https://dl.acm.org/doi/10.1145/3412841.3442057>. Acesso em: 26 jul. 2021.

WANG, Xiaoguang; CHANG, Wanli; TAN, Xu. Representing and Linking Dunhuang Cultural Heritage Information Resources Using Knowledge Graph. **KNOWLEDGE ORGANIZATION**, v. 47, n. 7, p. 604–615, 2020.

WIKIDATA. **Wikidata**. Disponível em: https://www.wikidata.org/wiki/Wikidata:Main_Page. Acesso em: 30 jul. 2021.

WIKIMEDIA COMMONS. **Wikimedia Commons**. Disponível em: https://commons.wikimedia.org/wiki/Main_Page. Acesso em: 30 jul. 2021.

WIKIMEDIA FOUNDATION. **List of Wikipedias - Meta**. Disponível em: https://meta.wikimedia.org/wiki/List_of_Wikipedias. Acesso em: 30 jul. 2021.

WIKIPÉDIA. Main Page. *In: Wikipedia, the free encyclopedia*. [s.l.: s.n.], 2021. Disponível em: https://en.wikipedia.org/w/index.php?title=Main_Page&oldid=1004593520. Acesso em: 30 jul. 2021.

ZENG, Marcia Lei. Interoperability. **Knowledge Organization**, v. 46, n. 2, 2019. Disponível em: <https://www.isko.org/cyclo/interoperability>. Acesso em: 30 jul. 2021.

