

**O PATRIMÔNIO CULTURAL NA WEB SEMÂNTICA***CULTURAL HERITAGE IN THE SEMANTIC WEB*

Juliana Martins de Castro Barroso¹
Universidade Federal de Minas Gerais

Célia da Consolação Dias²
Universidade Federal de Minas Gerais

RESUMO

A publicação de dados sobre o patrimônio cultural na web vem se ampliando a cada dia, e se adotados padrões e protocolos apropriados para a web semântica, tal movimento colabora para uma maior visibilidade do patrimônio cultural, possibilitando novas descobertas e análise de grandes conjuntos de dados. Existem diversos caminhos possíveis nesse sentido, e, assim, o objetivo deste estudo foi realizar um levantamento das iniciativas de publicação de dados de forma semântica na web, coletando relatos de pesquisas teóricas e aplicadas, nacionais e internacionais, com uma revisão integrativa de literatura. Diversas abordagens foram identificadas e divididas nas seguintes categorias: metadados, relações semânticas, agregação e interoperabilidade. As possibilidades identificadas constituem um material de grande relevância para orientar a elaboração de novas propostas no contexto nacional, ponto em que reside a principal relevância deste trabalho.

Palavras-chave: Patrimônio cultural; Web semântica; Metadados; Agregação; Interoperabilidade.

Keywords: Cultural heritage; Semantic web; Metadata; Aggregation; Interoperability.

1 INTRODUÇÃO

Instituições como galerias, bibliotecas, arquivos e museus, ou GLAMs (sigla para os termos em inglês: *galleries, libraries, archives and museums*), são responsáveis pela preservação e difusão de uma parcela considerável do patrimônio cultural nacional, e a internet tem se configurado como um dos principais meios

¹ Doutoranda pelo Programa de Pós-Graduação em Gestão e Organização do Conhecimento da Escola de Ciência da Informação da Universidade Federal de Minas Gerais. Orcid: <https://orcid.org/0000-0001-7848-0399>. E-mail: julianamartins@eci.ufmg.br.

² Docente do Programa de Pós-Graduação em Gestão e Organização do Conhecimento da Escola de Ciência da Informação da Universidade Federal de Minas Gerais. Orcid: <https://orcid.org/0000-0003-0891-6454>. E-mail: celiadias@eci.ufmg.br.

disponíveis para a promoção de acesso a esses bens. Percebe-se ainda que os projetos de publicação de dados sobre bens do patrimônio cultural fazem um uso cada vez maior dos princípios da web semântica, iniciativa que busca viabilizar a compreensão dos dados publicados pelas máquinas, a fim de garantir uma maior descoberta das informações sobre os bens, livre acesso a essas informações e possibilidade de integração de grandes conjuntos de dados.

Essas experiências mostram que o conhecimento dos princípios da web semântica, das ferramentas disponíveis e dos padrões de metadados adotados pelas comunidades interessadas é decisivo para iniciar o processo. Assim, o objetivo deste estudo foi reunir casos de publicação de acervos do patrimônio cultural na web semântica, buscando identificar metodologias e procedimentos que possam ser utilizados ou adaptados ao contexto das instituições nacionais.

2 FUNDAMENTAÇÃO TEÓRICA

O projeto da *World Wide Web*, ou apenas web, como ficou conhecida, tem evoluído desde sua origem em 1989, passando de uma “web de documentos” legíveis apenas por humanos para uma “web de dados” compreensíveis por máquinas. O foco da proposta da web semântica é capacitar os sistemas a compreenderem dados e suas relações. Nessa visão, os computadores “(...) teriam acesso a coleções de informações estruturadas e conjuntos de regras de inferência que poderiam ser utilizadas para conduzir um raciocínio automatizado” (BERNERS-LEE, HENDLER, LASSILA, 2001).

Para realizar essa transição, Souza e Alvarenga (2004, p. 134) destacam que “(..) é necessária uma padronização de tecnologias, de linguagens e de metadados descritivos, de forma que todos os usuários da web obedeçam a determinadas regras comuns e compartilhadas sobre como armazenar dados e descrever a informação armazenada”. Dessa forma, buscou-se na literatura estudos que apresentassem iniciativas de publicação de dados do patrimônio cultural fundamentados nos princípios da web semântica.

3 METODOLOGIA

Foi realizada uma revisão integrativa de literatura, metodologia que busca sintetizar a produção científica sobre um determinado tema (BOTELHO, CUNHA, MACEDO, 2011). As palavras-chave utilizadas na busca foram: “patrimônio cultural”, “web semântica”, “dados abertos conectados” e “interoperabilidade”, em português, inglês e espanhol. A busca foi realizada em sete bases de dados³. O recorte temporal foi definido entre 2010 e 2022, sendo a busca realizada em janeiro de 2022. Foram considerados artigos, *papers*, comunicações e capítulos de livros, revisados por pares. O *corpus* documental final foi composto por 173 estudos.

4 RESULTADOS PARCIAIS

Foram identificados quatro tipos principais de questões abordadas no *corpus* documental: uso de metadados no contexto da web semântica (112 estudos), relações semânticas entre conjuntos de dados (43 estudos), agregadores de conteúdo (27 estudos) e interoperabilidade (26 estudos).

O primeiro grupo temático consiste em estudos que identificam os padrões de metadados utilizados para publicar informações sobre bens do patrimônio cultural em projetos voltados para a web semântica. Foram identificados 22 esquemas de metadados nos estudos analisados, enfatizando sempre a importância do uso desses esquemas para estruturar os dados e facilitar a integração das informações publicadas posteriormente.

O segundo grupo é formado por trabalhos que analisam as relações semânticas entre conjuntos de dados, especialmente no âmbito de projetos de dados abertos conectados (objetivo principal da web semântica). São discutidas estratégias de enriquecimento semântico, anotações semânticas e desenvolvimento de ferramentas próprias.

Os agregadores de conteúdo constituem o terceiro tema abordado pela literatura. São ambientes que funcionam como pontos de curadoria digital sobre um

³ Biblioteca Digital Brasileira de Teses e Dissertações (BDTD), *Networked Digital Library of Theses and Dissertations* (NDLTD), *Web of Science* (WoS), Scopus – Elsevier, Base de Dados Referenciais de Artigos de Periódicos em Ciência da Informação (Brapci), Base de dados da revista *Knowledge Organization* – ISKO e Google Acadêmico.

assunto, como a Europeana, biblioteca digital da União Europeia considerada um modelo de sucesso na integração de dados de diferentes bens patrimoniais.

Por fim, a interoperabilidade é a temática predominante no quarto grupo de trabalhos, que se dedica ao mapeamento entre esquemas de metadados, seja para sua utilização em plataformas agregadoras, seja para sua conversão em conjuntos de dados conectados. Dentre as possibilidades de interoperabilidade citadas, destaca-se o *crosswalking*, “(...) um mapeamento de elementos, semântica e sintaxe de um esquema de metadados para outro” (CHAN, ZENG, 2006, p. 6).

5 CONSIDERAÇÕES FINAIS

A revisão de literatura apontou, por meio dos diferentes caminhos metodológicos apontados na seção anterior, as vantagens das tecnologias da web semântica na integração de dados, na pesquisa e recuperação de informações e na descoberta de novos *insights* com base na análise de grandes conjuntos de dados. A publicação de dados conectados resulta não apenas em maior visibilidade para os acervos culturais, mas também na ampliação das possibilidades de pesquisa, fruição, reuso e combinação de dados. Tanto as instituições quanto a sociedade têm muito a ganhar com a incorporação dos dados relacionados ao patrimônio cultural nacional na web semântica, e o primeiro passo nessa direção é o conhecimento desse campo de atuação. Prosseguindo com esta pesquisa, será investigado o panorama recente da publicação de dados pelas instituições patrimoniais nacionais, a fim de definir os próximos passos para a disponibilização de grandes conjuntos de dados compatíveis com as tecnologias semânticas atuais.

REFERÊNCIAS

BERNERS-LEE, T.; HENDLER, J.; LASSILA, O. The semantic web. **Scientific American**, v. 284, n. 5, p. 34-43, 2001.

BOTELHO, L. L. R.; CUNHA, C. C. A.; MACEDO, M. O método da revisão integrativa nos estudos organizacionais. **Gestão e sociedade**, v. 5, n. 11, p. 121-136, 2011.

CHAN, L. M.; ZENG, M. L. Metadata interoperability and standardization: a study of methodology - Part I. **D-Lib Magazine**, v. 12, n. 6, p. 1082-9873, 2006.

SOUZA, R. R.; ALVARENGA, L. A Web Semântica e suas contribuições para a ciência da informação. **Ciência da Informação**, v. 33, p. 132-141, 2004.