

Mapeamento de metadados no Tainacan: nova funcionalidade para os museus digitais vinculados ao Instituto Brasileiro de Museus

Metadata mapping in Tainacan: new functionality for digital museums linked to the Brazilian Institute of Museums

Joyce Siqueira

Universidade de Brasília, Brasil.

E-mail: joycitta@gmail.com

ORCID: 0000-0003-0246-8443

Dalton Lopes Martins

Universidade de Brasília, Brasil.

E-mail: daltonmartins@unb.br

Vinícius Nunes Medeiros

Tainacan.org, Brasil.

E-mail: viniciusmedeiros@inf.ufg.br

RESUMO

O Instituto Brasileiro de Museus – Ibram, em parceria com as Universidades Federal de Goiás e de Brasília vem construindo um novo serviço de busca e recuperação de objetos digitais dos acervos museológicos federais, um agregador de dados culturais intitulado Brasiliana Museum. O processo para agregação dos dados museológicos inclui diferentes etapas, sendo uma delas, essencial, o mapeamento dos metadados dos museus de origem a um padrão de metadados único, viabilizando a interoperabilidade de informação entre os museus. Os museus do Ibram disponibilizam seus acervos por meio do software Tainacan, dessa forma, este artigo apresenta uma nova funcionalidade para o Tainacan, intitulada “Tainacan Mappers IBRAM

Como citar: Siqueira, J.; Martins, D. L.; & Medeiros, V. N. (2022). Mapeamento de metadados no Tainacan: nova funcionalidade para os museus digitais vinculados ao Instituto Brasileiro de Museus. En T.M.R. Dias (Ed.), Informação, Dados e Tecnologia. *Advanced Notes in Information Science, volume 2* (pp. 182-191). Tallinn, Estonia: ColNes Publishing. DOI: 10.47909/anis.978-9916-9760-3-6.92.

Copyright: © 2022, The author(s). This is an open access work distributed under the terms of the CC BY-NC 4.0 license which permits copying and redistributing the material in any medium or format, adapting, transforming and building upon the material as long as the license terms are followed.

INBCM”, que visa permitir aos usuários mapear seus acervos para os elementos descritivos propostos pelo Ibram, dispostos no Inventário Nacional dos Bens Culturais Musealizados – INBCM. A pesquisa é de natureza qualitativa, de cunho exploratório e descritivo, na modalidade de estudo de caso. Os resultados apresentam o *script* desenvolvido utilizando a linguagem de programação *Hipertext Preprocessor - PHP*, e os benefícios do mesmo para à agregação dos dados. A nova funcionalidade está em uso, no protótipo do *Brasiliana Museum*, e como trabalhos futuros, pretende-se implementar operações mais avançadas, tais como, a transformação, a concatenação e a divisão de valores.

Palavras-chave: Tainacan; Brasiliana Museum; Agregação; Mapeamento; Interoperabilidade.

ABSTRACT

The Brazilian Institute of Museums – Ibram, in partnership with the Federal University of Goiás and the University of Brasília, has been building a new service for searching and retrieving digital objects from federal museum collections, an aggregator of cultural data called Brasiliana Museums. The process for aggregating museological data includes different steps. One of which is essential is mapping metadata from the museums of origin to a single metadata standard, enabling the interoperability of information between museums. Ibram’s museums make their collections available through the Tainacan software. Thus, this article presents a new functionality for Tainacan, entitled “Tainacan Mappers IBRAM INBCM”, which aims to allow users to map their collections to the descriptive elements proposed by Ibram, listed in the National Inventory of Museum Cultural Assets – INBCM. The research is qualitative, exploratory and descriptive, in the form of a case study. The results presented the script developed using the programming language Hypertext Preprocessor - PHP and its benefits for data aggregation. The new functionality is in use in the Brasiliana Museum prototype. As future works, it is intended to implement more advanced operations, such as transformation, concatenation and division of values.

Keywords: Tainacan; Brasiliana Museums; Aggregation; Crosswalk; Interoperability.

INTRODUÇÃO

O INSTITUTO Brasileiro de Museus – Ibram, em parceria com as Universidades Federal de Goiás e de Brasília está desenvolvendo o *Brasiliana Museum*, um agregador de objetos digitais culturais dos acervos dos museus vinculados ao Ibram, utilizando o Tainacan,¹ um *software* para a criação e gestão de coleções digitais, tanto nos museus de origem quanto no agregador.

¹ <https://tainacan.org/>

Com esse intento, é necessário que todos os repositórios sejam interoperáveis. Por padrão, o Tainacan oferece *Universal Resource Locators* - URLs alternativos que possibilitam com que outras aplicações possam se beneficiar dos dados presentes no repositório. No entanto, coletar e apenas reunir os dados dos museus, tais como estão na origem, não é suficiente. É essencial assegurar que estes dialoguem entre si, ou seja, garantir que possuam o mesmo padrão de metadados, resultando em uma mesma abordagem sintática e semântica na agregação (NISO, 2004; Zeng & Qin, 2016).

Na literatura, há diferentes formas de se assegurar este diálogo. Nesta pesquisa, utilizaremos o processo denominado ‘mapeamento’ ou ‘crosswalk’, quando elementos de um padrão de metadados são associados a outros, viabilizando a interoperabilidade e permitindo que coleções heterogêneas possam ser pesquisadas simultaneamente (Chan & Zeng, 2006). Pretende-se testar e comparar os resultados com outros métodos em estudos futuros.

No caso do Ibram, em 2014, foram criados elementos descritivos que devem ser utilizados por todos os museus federais, intitulado ‘Inventário Nacional dos Bens Culturais Musealizados’ – INBCM, composto por 15 elementos de metadados (Brasil, 2014). No entanto, embora de utilização obrigatória por todos os museus vinculados ao Ibram, a realidade nos apresenta museus que utilizam esses elementos de forma parcial, não utilizando todos os metadados requisitados no modelo e/ou utilizando metadados que não constam no INBCM. Sendo assim, torna-se essencial a etapa de mapeamento dos dados, mesmo que haja uma determinação de uso do INBCM.

Para oferecer uma solução completa ao Ibram, foi desenvolvido uma nova funcionalidade, um *plugin*, que habilita no Tainacan a interface de mapeamento nas instalações (museus que serão agregados), resultando, ao final, em URL’s para arquivos no formato *JavaScript Object Notation* - JSON, mapeados de acordo com o modelo do INBCM, e prontos para serem coletados pelo Brasileira Museum. Este artigo apresenta os detalhes do resultado do desenvolvimento e utilização deste *plugin* pelos museus vinculados ao Ibram.

OBJETIVO

Apresentar uma nova funcionalidade desenvolvida para o *software* Tainacan, um *plugin* para mapeamento de metadados para os elementos descritivos do INBCM, com intuito de viabilizar a interoperabilidade de informações entre os acervos museológicos digitais do Ibram, visando ofertar um novo serviço de busca e recuperação de objetos digitais dos acervos

museológicos federais, um agregador de dados culturais intitulado Brasileira Museums.

PROCEDIMENTOS METODOLÓGICOS

Pesquisa de natureza qualitativa, de cunho exploratório e descritivo, que aborda o processo de mapeamento dos metadados dos acervos museológicos digitais do Ibram, configurando a pesquisa na modalidade de estudo de caso. Como procedimento técnico, utilizou-se de pesquisa bibliográfica no intuito de fornecer sustentabilidade teórica ao estudo. O *plugin* intitulado “Tainacan Mappers IBRAM INBCM” foi desenvolvido utilizando a linguagem de programação *Hypertext Preprocessor* - PHP,² e está disponível no GitHub,³ sob a licença GPLv3,⁴ para uso no *software* Tainacan, versão 0.15 ou superior, de código aberto, gratuito, também sob a licença GPLv3.

FUNDAMENTOS TEÓRICOS

O Brasileira Museums tem como objetivo agregar os acervos digitais dos museus vinculados ao Ibram e, de modo que tal proposição possa se tornar viável, questões relacionadas à interoperabilidade precisam ser fundamentadas em seus aspectos teóricos-metodológicos, conforme se apresentam a seguir. Niso (2004) conceitua interoperabilidade como a capacidade de múltiplos sistemas, com diferentes plataformas de *hardware* e *software*, com diferentes estruturas de dados ou com diferentes *interfaces*, trocar dados com perda mínima de conteúdo e funcionalidade, de modo a garantir a troca de informações de maneira eficaz e eficiente.

Segundo Arms (2002) citado por Fusco (2010), interoperar requer acordos de cooperação em três níveis: técnico, de conteúdo e organizacional. O nível técnico diz respeito a interoperabilidade tecnológica, como exemplo, os protocolos e formatos de intercâmbio. O nível de conteúdo, trata sobre a interoperabilidade semântica, garantindo que os dados trocados tenham seu significado corretamente interpretado. O nível organizacional, trata a interoperabilidade política, quando organizações se reúnem com intuito de obter a interoperabilidade.

Nesta pesquisa, o foco está na interoperabilidade de conteúdo ou da informação. Para Fusco (2010), do ponto de vista da interoperabilidade de conteúdo é necessária a compreensão de três níveis: semântico, estrutural e sintático. No semântico, deve-se entender o significado de cada

² <https://www.php.net/>

³ <https://github.com/tainacan/tainacan-mappers-ibram-inbcm>

⁴ <https://www.gnu.org/licenses/gpl-3.0.pt-br.html>

elemento e suas associações. No estrutural, cada elemento de um padrão de metadados, descreve os seus tipos, a escala de valores possíveis e os mecanismos utilizados para relacionar esses elementos. No sintático, são definidos como os metadados devem ser codificados para a transferência de informações.

Para Chan e Zeng (2006), idealmente, uma abordagem padrão uniforme garantiria a interoperabilidade máxima, por exemplo, se todos fossem obrigados a usar o Dublin Core, um alto nível de consistência seria mantido. No entanto, embora seja uma solução conceitualmente simples, nem sempre é viável na prática. Considerando que haverá inúmeros padrões de metadados, garantir a interoperabilidade requer o uso, individual ou concomitante, de diferentes métodos, dos quais Chan e Zeng (2006) citam: derivação, perfis de aplicação, mapeamento, comutação, estrutura e registro. Nesta pesquisa, detalharemos o mapeamento ou *crosswalk*.

Segundo Chan e Zeng (2006) mapear um padrão de metadados para outro é o método mais frequentemente utilizado para interoperar. Para tal, algumas propriedades de metadados precisam ser consideradas para o mapeamento e para elencá-las citam Pierre e Laplant (1998): um identificador único para cada elemento de metadados; uma definição semântica de cada elemento de metadados; identificar se um elemento de metadados é obrigatório, opcional ou obrigatório com base em condições; identificar se um elemento de metadados pode ou não ocorrer várias vezes; identificar as relações hierárquicas pai-filho; identificar restrições impostas ao valor do elemento (por exemplo, texto livre, intervalo numérico, data ou um vocabulário controlado) e; identificar necessidade de suporte opcional para elementos de metadados definidos localmente.

Chan e Zeng (2006) destacam que há duas abordagens de mapeamento utilizadas na prática, a absoluta e a relativa. No caso do *plugin* do Ibram, utilizou-se o mapeamento absoluto, que requer mapeamento exato entre os elementos envolvidos de um esquema de origem e um esquema de destino, de forma que onde não há equivalência exata, não há mapeamento. O mapeamento absoluto garante a equivalência dos elementos, porém, os metadados não mapeados são descartados, ocorrendo em grande perda, especialmente quando o esquema de origem possui uma estrutura mais rica do que a do esquema de destino. Outro ponto relevante, é a ocorrência de diferentes graus de equivalência: um para um, um para muitos, muitos para um e um para nenhum. Por exemplo, em um para um, o metadado de origem encontra equivalência em um único metadado de destino (autor – autor); em muitos para um, mais de um metadado de origem é necessário para equivalência do metadado de destino (matérias e técnicas – material/técnica).

Normalmente apenas os nomes dos elementos e suas definições são levados em consideração no mapeamento. Porém, Pierre e Laplant (1998) ressaltam que para que haja o mapeamento faz necessária a realização de uma tarefa intelectual, que deve ser desenvolvida por especialistas da área, para determinar o mapeamento semântico dos elementos, ou seja, especificar qual elemento da origem deve ser relacionado semanticamente ao elemento de destino. Chan e Zeng (2006) ao tratarem sobre o mapeamento, usam o termo padrão de metadados, que pode ser definido como uma especificação aberta que relaciona um conjunto de elementos e seus significados, o que promove um vocabulário comum capaz de descrever uma variedade de estruturas de dados no sentido de atender a várias comunidades. Além disso, cada elemento possui um conjunto de regras que especifica o valor de conteúdo permitindo sua relação com outros elementos, e cuja composição se dá por um vasto leque de formalidades para tratamento nos dados.

Em sua concepção, geralmente há um esforço conjunto entre comunidades e entidades normativas interessadas em fornecer um *framework* comum de metadados para aplicações de interesses diversos, como nos seguintes casos: *International Organization for Standardization - ISO*, *American National Standards Institute - ANSI*, *Dublin Core Metadata Initiative - DCMI*, *Library of Congress - LC*, *Getty Research*, *Wide Web Consortium (W3C)*, entre outros (Pierre & Laplant, 1998; Lima *et al.*, 2016). Há uma grande variedade de padrões de metadados, alguns com pouco rigor formal (Allemang *et al.*, 2020), os quais são usados em diversos domínios de assunto. Na área do patrimônio cultural, citamos alguns, tais como o *Categories for the Description of Works of Art - CDWA*, o *Cataloging Cultural Objects - CCO*, *CDWA Lite* e *LIDO*, o *VRA Core* e o *Modelo Conceitual de Referência (CRM)*, do *CIDOC*.

Embora sejam padrões de metadados internacionais e bastante utilizados, todos são, em diferentes níveis, complexos para uso. No caso dos museus federais, o Ibram optou por não utilizar esses padrões, mas por definir regras específicas para catalogação de seus acervos. O resultado é a Resolução Normativa nº 2, de 29 de agosto de 2014, que estabelece os elementos de descrição das informações sobre o acervo museológico, bibliográfico e arquivístico que devem ser declarados no Inventário Nacional dos Bens Culturais Musealizados, o INBCM.

O INBCM possui 15 metadados: número de registro, outros números, situação, denominação, título, autor, classificação, resumo descritivo, dimensões, material/técnica, estado de conservação, local de produção, data de produção, condições de reprodução e mídias relacionadas. O documento apresenta a descrição semântica de cada elemento, além da informação de sua obrigatoriedade ou não (Brasil, 2014).

TRABALHOS RELACIONADOS

Pramudyo e Hendrawan (2020) apresentam estudo que visa compreender, descrever e analisar a interoperabilidade de metadados na Biblioteca Universitas Brawijaya que usou o *software* Brawijaya Knowledge Garden e Eprints, na Biblioteca da Universidade de Muhammadiyah Malang que usou a Biblioteca Digital Ganesha e o *software* Eprints, e na Malang Biblioteca estadual que usava o *software* Muatan Lokal. A descoberta indica que a interoperabilidade de metadados pode ser realizada usando mapeamento de metadados. Barroso, Hartimann e Ribeiro (2015) discutem a agregação dos dados da Biblioteca Digital de Arte (BDArt) com os documentos da biblioteca e do acervo da Escola de Belas Artes da Universidade do Porto. A interoperabilidade entre as coleções do arquivo, da biblioteca e do museu é necessária porque muitas obras alocadas a diferentes coleções estão intimamente relacionadas e só podem ser vistas como um todo por meio de funcionalidades de mapeamento das coleções.

Yamamoto *et al.* (2013) propõem um novo conjunto de metadados, intitulado *Museus Core*, que pode ser mapeado mecanicamente a partir dos bancos de dados originais. No artigo, os autores também discutem o mapeamento do *Museus Core* para *Dublin Core* e a utilidade do novo conjunto de elementos. Os trabalhos apresentam resultados satisfatórios na interoperabilidade entre os sistemas com a realização do mapeamento dos dados, sendo um indicativo importante para realização desta pesquisa. Além disso, Yamamoto, Adachi e Hachimura (2013) propõe um novo conjunto de metadados, assim como o Ibram desenvolveu os elementos descritivos, o INBCM, validando a metodologia empregada.

4. RESULTADOS

O *plugin* de mapeamento está disponível para instalação no WordPress e pode ser adicionado a versão 0.15 ou superior do Tainacan. Após sua instalação, oferece uma interface simplificada, que visa mapear os metadados de uma coleção, conforme Figura 1, na qual os itens do INCBM estão dispostos à esquerda e podem ser relacionados aos metadados da coleção.

O mapeamento é feito no grau de equivalência um para um. Nos elementos que não há equivalência exata os metadados são desconsiderados. No Tainacan, os usuários são livres para modelar quaisquer metadados que desejarem e determinar se eles são públicos ou privados. Dada essa característica, embora o INBCM descreva quais metadados são obrigatórios ou opcionais e não especifique quais devem ser mantidos privados ou públicos, o mapeador considera todos os metadados existentes, desprezando as validações de metadados privados e obrigatórios. Porém, só

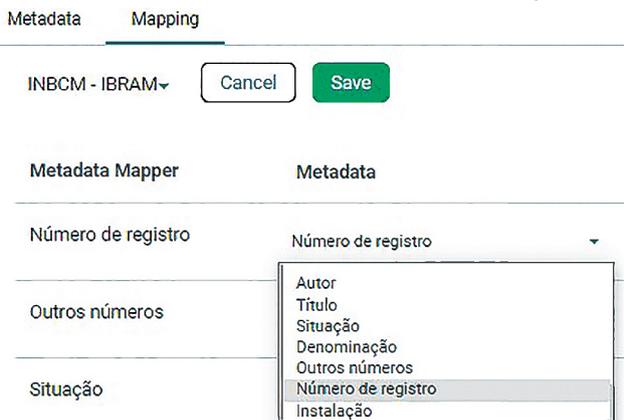


Figura 1. Interface de mapeamento dos metadados para INCBM, no Tainacan.

Fonte: Elaborada pelos autores, 2021.

disponibiliza no arquivo final, no formato JSON, os metadados mapeados e públicos. Para exemplificar, a Tabela 1 apresenta parte do mapeamento do Museu da Inconfidência, que possui, ao todo, 50 metadados, sendo 23 privados e 27 públicos.

INBCM	MUSEU DA INCONFIDÊNCIA		MAPEADO?
Metadados	Metadados	Status	
N. de registro	N. de registro	público	Sim
Outros números	Outros números	privado	Sim
Diâmetro	Não há	público	Não
Não há	Editor	privado	Não

Tabela 1. Mapeamento parcial do Museu da Inconfidência.

Fonte: Elaborado pelos autores, 2021.

Assim, conforme a Tabela 1, embora seja possível mapear dados privados, apenas os públicos serão disponibilizados, não havendo necessidade do usuário se preocupar com essas questões. Caso haja um item no INBCM que não haja representação nos metadados de origem, este não é mapeado e é descartado, o mesmo ocorre quando não há representação no INBCM, independentemente de serem privados ou públicos.

O *plugin* ainda está sendo aperfeiçoado e está previsto para as próximas etapas a implementação do mapeamento muitos para um, por exemplo, o metadado Material/Técnica traz dois conteúdos distintos, que em alguma museus são separados em dois metadados separados, Matérias e

Técnicas, sendo necessário reunir os dois em um único. Um para muitos ainda não foi detectada a necessidade, considerando o caso Ibram. Em relação ao mapeamento semântico, propriamente dito, ele é realizado por pessoal especializado.

5. CONSIDERAÇÕES FINAIS

As pesquisas relativas à construção do serviço de agregação de acervos digitais de museus, o Brasiliana Museus, se desdobram em outras, necessárias a efetivação da agregação. Uma delas, o mapeamento, embora simples, é essencial. Por ser um plugin, também serve como inspiração para sejam desenvolvidos novos plugins, para outros repositórios, visando a interoperabilidade entre sistemas.

Como trabalhos futuros, uma nova versão deve implementar novas possibilidades de se mapear os metadados dos acervos, possibilitando que operações mais avançadas possam ser realizadas, por exemplo, a transformação, concatenação e divisão de valores presentes no repositório.

CONFLITOS DE INTERESSE

Os autores declaram não haver conflito de interesses.

DECLARAÇÃO DE CONSENTIMENTO DE DADOS

O código do plugin de mapeamento de dados para o INBCM desenvolvido para o Ibram, gerado durante o desenvolvimento deste estudo foi depositado no GitHub e está acessível em <https://github.com/tainacan/tainacan-mappers-ibram-inbcm>

FINANCIAMENTO

Bolsa de doutorado recebida pelo CNPq. 

REFERÊNCIAS

- ALLEMANG, D., HENDLER, J., & GANDON, F. (2020). *Semantic Web for the Working Ontologist: Effective Modeling for Linked Data, RDFS, and OWL* (3a ed.). Association for Computing Machinery.
- BARROSO, I, HARTMANN, N & RIBEIRO, C. (2015). Metadata Crosswalk for a Museum Collection in a Thematic Digital Library. *Journal of Library Metadata*, 15, 36–49. <https://doi.org/10.1080/19386389.2015.1011025>

- BRASIL. Instituto Brasileiro de Museus, Resolução Normativa nº 2 (2014) Diário Oficial. Recuperado de <https://pesquisa.in.gov.br/imprensa/jsp/visualiza/index.jsp?jornal=1&pagina=14&data=01/09/2014>
- CHAN, L. M.; & ZENG, M. L. (2006). Metadata Interoperability and Standardization. A Study of Methodology Part I. Archiving Interoperability at the Schema Level. *D-Lib Magazine*, 12(6). <https://doi.org/10.1045/june2006-chan>
- FUSCO, E. (2010). *Modelos conceituais de dados como parte do processo de catalogação: perspectiva do uso dos FRBR no desenvolvimento de catálogos bibliográficos digitais* (Tese de doutorado). Universidade Estadual Paulista Júlio de Mesquita Filho, Brasil.
- LIMA, F. R. B., VAC SANTOS, P. L., & SANTAREN, J. E. (2016). Padrão de metadados no domínio museológico. *Perspectivas em Ciência da Informação*, 21, 50-69. <https://doi.org/10.1590/1981-5344/2639>
- MARCONDES, C. H. (2016). Interoperabilidade entre acervos digitais de arquivos, bibliotecas e museus: potencialidades das tecnologias de dados abertos interligados. *Perspectivas em Ciência da Informação*, 21, 61-83. <https://doi.org/10.1590/1981-5344/2735>
- NISO (National Information Standards Organization), (2004). *Understanding Metadata: What is Metadata, and What is it For?: A Primer*. Bethesda, MD: NISO Press.
- PIERRE, M. ST. & LAPLANT, W. P. (1998). *Issues in crosswalking content metadata standards*. Baltimore, Maryland, USA: NISO. https://groups.niso.org/publications/white_papers/crosswalk/
- PRAMUDYO G. N., HENDRAWAN M. R. (2020) Metadata Interoperability for Institutional Repositories: A Case Study in Malang City Academic Libraries. In: Ishita E., Pang N.L.S., Zhou L. (eds), Digital Libraries at Times of Massive Societal Transition. ICADL 2020. *Lecture Notes in Computer Science*, vol 12504. Springer, Cham. https://doi-org/10.1007/978-3-030-64452-9_33
- YAMAMOTO, Y., ADACHI, F., & HACHIMURA, K. (2013, September). Common Metadata to Search for Non-digital Cultural Resources in Heterogeneous Databases. In *2013 International Conference on Culture and Computing* (pp. 224-225). IEEE. <https://doi.org/10.1109/CultureComputing.2013.76>
- ZENG, M. L. & QIN, J. (2016). *Metadata Standards*. In: _____. *Metadata* (2ª Ed.). Chicago: Neal-Schuman/American Library Association.

