TPBCI | Tendências da Pesquisa Brasileira em Ciência da Informação

ACERVOS AGREGADOS DO INSTITUTO BRASILEIRO DE MUSEUS: DESENVOLVIMENTO DO PAINEL DE VISUALIZAÇÃO ANALÍTICA[1]

BRAZILIAN INSTITUTE OF MUSEUMS'S AGGREGATE COLLECTIONS: DASHBOARD DEVELOPMENT

Joyce Siqueira [2]
Dalton Martins Lopes [3]

Resumo: Na web há, atualmente, 19 museus vinculados ao Instituto Brasileiro de Museus – Ibram, que disponibilizaram seus acervos digitais por meio do *software* Tainacan. A partir deste cenário, deu-se início a um novo estudo em vistas a desenvolver um agregador dos acervos do Ibram. A partir da implementação do protótipo do agregador, que utiliza como tecnologia a pilha Elastic Stack, composta pelas ferramentas Beats, Logstash, Elasticsearch e Kibana, observou-se que também seria possível, por meio do Kibana, criar painéis dos acervos digitais integrados do Ibram. Os painéis possibilitam a visualização analítica, por meio de interfaces gráficas, impulsionada pelo crescimento das bases dados e dos processos de gestão de conhecimento. Dessa forma, este artigo descreve o processo de criação do painel para o Ibram. Trata-se de uma pesquisa qualitativa, exploratória e descritiva, que, envolvendo o processo de agregação dos acervos do Ibram, configura-se na modalidade de estudo de caso. O resultado é um painel interativo, atualizado automaticamente, composto por 19 museus, totalizando 16.057 itens, que oferece ao Ibram recursos para uma gestão da documentação museológica mais eficaz, pois apresenta uma visão macro dos museus. Além disso, dados integrados podem demonstrar inconsistências sintáticas ou semânticas, possibilitando aos museus melhorarem seus dados na origem.

Palavras-Chave: Agregação de dados. Ferramenta Kibana. Museus. Painel de Visualização. *Software* Tainacan.

Abstract: *On the web there are currently 19 museums linked to the Brazilian Institute of Museums – Ibram, which have made their digital collections available through the Tainacan software. From this scenario, a new study was started with a view to developing an aggregator of Ibram's collections. From the implementation of the aggregator prototype, which uses the*

[1] Este trabalho foi submetido, avaliado, aprovado, apresentado e premiado no GT 8 - Informação e Tecnologia, na modalidade Trabalho Completo, do XXI ENANCIB – 2021.

[2] Doutoranda em Ciência da Informação. Universidade de Brasília. joycitta@gmail.com. https://orcid.org/0000-0003-0246-8443.

[3] Doutor em Ciência da Informação. Universidade de Brasília. daltonmartins@unb.br. https://orcid.org/0000-0002-6244-6791.

Copyright: Esta obra está licenciada com uma Licença Creative Commons Atribuição 4.0 Internacional. (cc) BY

Elastic Stack as technology, composed by the Beats, Logstash, Elasticsearch and Kibana tools, it was observed that it would also be possible, through Kibana, to create panels of Ibram's integrated digital collections. Dashboards enable analytical visualization through graphical interfaces, driven by the growth of databases and knowledge management processes. Thus, this article describes the process of creating the dashboard for Ibram. This is a qualitative, exploratory and descriptive research, which, involving the aggregation process of Ibram's collections, is configured as a case study. The result is an interactive panel, automatically updated, composed of 19 museums, totaling 16,057 items, which offers Ibram resources for a more effective management of museological documentation, as it presents a macro view of the museums. Furthermore, integrated data can demonstrate syntactic or semantic inconsistencies, enabling museums to improve their data at source.

Keywords: *Dashboard. Data aggregation. Kibana Tool. Museums. Tainacan software.*

1 INTRODUÇÃO

Novas tecnologias propiciaram a reinvenção do espaço museal, dando às instituições culturais a oportunidade de extrapolarem a missão de custodiar, preservar e manter os acervos, as permitindo aperfeiçoar a comunicação com o público em geral, disseminando seu conteúdo para que se constitua como recurso de memória, de história, de cultura, de educação e de cidadania (MARCONDES, 2016; MUCHACHO, 2005).

Neste contexto, a democratização do acesso por meio dos acervos digitais potencializa a propagação do conhecimento, além de ampliar exponencialmente a visibilidade das instituições, anteriormente restrita à visitação in loco, trazendo um novo fôlego para museus, arquivos e bibliotecas no século 21 (INSTITUTO BRASILEIRO DE MUSEUS, 2022; SANTAREM SEGUNDO; SILVA; MARTINS, 2019).

Instituições responsáveis pela guarda de acervos do patrimônio cultural, nacionais e, principalmente, internacionais, têm investido na digitalização e difusão de seus acervos, resultando em inúmeros museus digitais distribuídos na web.

Esse cenário, embora muito rico, trouxe desafios à descoberta dos recursos disponíveis. Por exemplo, para localizar um determinado objeto digital é preciso identificar o acervo a qual este pertence e ter conhecimentos prévios mínimos sobre o objeto, o que para um usuário leigo poderia ser um sério entrave, limitando, sobremaneira, a real difusão do acervo.

Destacamos que agregar dados não é algo novo, nem trivial, nem tão pouco exclusividade da área cultural. Diferentes iniciativas, de diferentes áreas, vêm, a algum tempo, desenvolvendo soluções que melhor se adequem às suas necessidades.

Há soluções nas áreas da Bioinformática, da Educação, da Geografia, da Medicina, da Música, entre outras. Por exemplo, na Bioinformática, integram-se dados de expressão gênica da *Drosophila melanogaster*, conhecida como mosca da fruta, para selecionar genes e validar resultados experimentais (MILES *et al*., 2010); na Geografia, integram-se dados de localizações geográficas para colaborar com a avaliação de pedidos de licenças para uso da água (ZIÉBELIM *et al*., 2017); na Medicina, integram-se dados de formulários e sistemas para estudos epidemiológicos para determinar as

causas de doenças, como obesidade, depressão, demência (KIRSTEN *et al*., 2017); na Música, integra-se semanticamente informações de diferentes músicas para oferecer serviços de qualidade aos usuários (GANGLER, 2011).

No entanto, agregar dados culturais constituem uma camada extra de dificuldade, pela diversidade de tipos de objetos digitais. Assim, não há uma solução simples, mas pelo contrário, há problemas bastante desafiadores de engenharia de *software* e modelagem de dados (AVAZPOUR; GRUNDY; ZHU, 2019).

Em se tratando do patrimônio cultural, Freire *et al*. (2017) realizaram um levantamento sobre as tecnologias web para agregação de dados, trazendo importantes constatações. Primeiro que, neste domínio, a abordagem tecnológica majoritária é o protocolo Open *Archives Initiative Protocol for Metadata Harvesting* (OAI-PMH), que mantém um bom grau de interoperabilidade, mas que com o avanço tecnológico deixou ter a motivação para sua a adoção tão clara quanto costumava ser; segundo, destacam especificidades que dificultam o processo de agregação, tais como: a heterogeneidade de subdomínios, composto por bibliotecas, arquivos e museus, que possuem suas próprias práticas e padrões, e os orçamentos limitados, que tornam mais lenta a velocidade e a extensão da inovação e adoção de novas tecnologias; terceiro, os metadados são transferidos para o agregador, mas continuam a evoluir no provedor de dados, portanto, o agregador precisa atualizar periodicamente sua cópia dos dados.

Outro fator crucial, também destacado por Freire *et al*. (2017), é a prática de acordar o uso de um único modelo de dados, que geralmente tem a mesma semântica dos dados originais e que atendam às necessidades de informação dos serviços prestados pelo agregador.

O cenário internacional apresenta diferentes instituições que ofertam aos usuários finais uma interface única para busca e recuperação de dados, por meio de uma plataforma centralizadora que direciona os acessos aos museus de origem. Diferentes realidades, mas que têm alterado a maneira de se coletar e disseminar conteúdo digital. Podemos citar, a *American Art Collaborative* (AAC), a *Digital Public Library of America* (DPLA), a *World Digital Library* e a *Google Arts & Culture*, dos Estados Unidos da América; a DigitalNZ, da Nova Zelândia; a Europeana, da Europa; a Mexicana, do México; a TROVE,

da Austrália (SIQUEIRA; MARTINS, 2019a, 2020, 2021b; SANTAREM SEGUNDO; SILVA; MARTINS, 2019).

Outro tópico importante, é o retorno financeiro que os agregadores internacionais trouxeram. Duas importantes instituições, a *British Library* e a Europeana, realizaram estudos, ambos publicados em 2013, que indicaram que os serviços on-line da *British Library* estão avaliados em 19.5 milhões de libras por ano, gerando benefícios a estudantes, pesquisadores e cidadãos do Reino Unido. A Europeana concluiu que o impacto econômico para o continente europeu seria de aproximadamente 78.8 milhões de euros em um cenário de base, 59.6 milhões de euros em um cenário pessimista e 97.7 milhões em um cenário otimista (MARTINS; SILVA; CARMO, 2018).

No Brasil, há algumas importantes iniciativas que objetivam integrar acervos culturais, das quais citamos três, a Brasiliana Iconográfica, que visa integrar fontes iconográficas sobre o Brasil, datadas do século XVI até as primeiras décadas do século XX (BRASILIANA ICONOGRAFICA, 2017); a Brasiliana Fotográfica, que visa integrar os acervos originais de documentos fotográficos referentes ao Brasil (BRASILIANA FOTOGRAFICA, 2022a) e; a Rede Web de Museus, que visa estabelecer uma política estadual integrada e colaborativa para os museus do Estado do Rio de Janeiro (GOVERNO DO RIO DE JANEIRO, 2022b). Todas abertas a expansão de seus serviços a outras instituições nacionais e internacionais, públicas ou privadas interessadas (BRASILIANA ICONOGRAFICA, 2017; BRASILIANA FOTOGRAFICA, 2022a; GOVERNO DO RIO DE JANEIRO, 2022b).

Porém, de forma geral, as iniciativas brasileiras são limitadas, isoladas e necessitam de apoio técnico e de modelos de governança escaláveis para o território brasileiro. Além disso, identifica-se a falta de incentivo político e econômico para ofertar soluções nessa direção (SANTARÉM SEGUNDO; SILVA; MARTINS, 2019). Há em curso algumas iniciativas no país visando desenvolver novas ações de agregação de acervos em plataformas digitais. Cabe destacar a iniciativa do projeto Plataforma Acervo, tema da presente pesquisa.

O Instituto Brasileiro de Museus (IBRAM), criado em 2009, é uma autarquia vinculada ao Ministério do Turismo, responsável pela Política Nacional de Museus e pela melhoria dos serviços do setor, tais como: o aumento de visitação e a arrecadação dos

museus; o fomento de políticas de aquisição; a preservação de acervos e; a criação de ações integradas entre os museus brasileiros. Além de administrar 30 instituições museológicas, distribuídas no Sul, Sudeste, Centro-Oeste e Nordeste do país (BRASIL, 2022).

Em 2014, o Ibram firmou parceria com o Laboratório de Políticas Públicas Participativas, da Universidade Federal de Goiás - UFG, parceria essa, atualmente transferida para a Universidade de Brasília - UnB, visava o desenvolvimento do Projeto Tainacan, com intuito de promover os fundamentos de uma política nacional para acervos digitais (BRASIL, 2021). Em 2016, a primeira versão do *software* Tainacan é apresentada.

> O Tainacan é uma ferramenta flexível e poderosa para WordPress que permite a gestão e a publicação de coleções digitais com a mesma facilidade de se publicar posts em blogs, mas mantendo todos os requisitos de uma plataforma profissional para repositórios (BRASIL, 2021).

O *software* foi pensado para ser uma solução tecnológica para a difusão e para a interoperabilidade de acervos digitais das instituições culturais brasileiras (MARTINS; CARVALHO JÚNIOR; GERMANI, 2019). É um *software* livre e gratuito, que vislumbrou sua manutenção a longo prazo, escolhendo a ecologia do WordPress, amplamente difundido e com uma comunidade de desenvolvedores e usuários ativa, além da facilidade, para que profissionais da área da cultura, sem formação específica em Tecnologia da Informação - TI, pudessem instalar, customizar e utilizar de forma autônoma (MARTINS; MARTINS, 2020).

Atualmente, o Tainacan se configura como uma consolidada ferramenta utilizada por diferentes instituições, incluindo Universidades Públicas (MARTINS; MARTINS, 2020), Museus Estaduais e Municipais, Instituições privadas, Instituições internacionais e pelo Ibram, que possui 19 museus disponíveis na web e outros em fase de implantação do Tainacan (TAINACAN, 2022; BRASIL, 2021).

Em uma nova etapa de implementação, o Laboratório de Inteligência de Redes da Universidade de Brasília (UnB) tem desenvolvido os primeiros serviços de integração das coleções museológicas publicadas pelo Ibram visando a criação de uma interface

única de busca como experimento dos efeitos da publicação de acervos em rede (BRASIL, 2021).

A partir de uma revisão sistemática de literatura, Siqueira e Martins (2021) levantaram as principais tecnologias utilizadas por projetos de agregação de dados culturais, além de buscarem na literatura cinzenta por novas tecnologias capazes de realizar a agregação de dados. Nesse sentido, o grupo de pesquisadores do Laboratório optou pelo uso do *software* Tainacan e da pilha Elastic Stack, que tem como principal vantagem a automatização de etapas importantes e a facilidade e agilidade nas buscas (ELASTIC, 2022b).

No entanto, durante o processo de prototipação da busca integrada, identificou-se a possibilidade de oferecer painéis de visualização dos acervos digitais integrados do Ibram. Segundo Sarikaya *et al*. (2019, p. 683, tradução nossa):

> Dados onipresentes e tecnologias de visualização disponíveis ao público ampliaram a adoção do painel para novos domínios. Consequentemente, o conceito de painel evoluiu de telas de relatório para incluir interfaces interativas, com múltiplas visualizações e propósitos, incluindo comunicação, aprendizado e motivação, além das noções clássicas de monitoramento e suporte à decisão.

Os painéis são extremamente importantes em um mundo baseado em dados. Empresas, organizações sem fins lucrativos e grupos comunitários dependem de painéis todos os dias para realizar seu trabalho. Para muitos, os painéis podem ser o primeiro (ou único) encontro com todos os dados. Além disso, o uso do painel ultrapassou os limites organizacionais, incluindo os indivíduos e o público em geral (SARIKAYA *et al*., 2019).

Os painéis são ferramentas para a visualização da informação, especificamente, nesta pesquisa, a 'Visualização Analítica', que procura focar no pensamento analítico por meio de interfaces gráficas, impulsionada pelo crescimento das bases dados e dos processos de gestão de conhecimento, bem como do processo de diminuição de incertezas para tomada de decisão e análise do comportamento cognitivo (FERREIRA, 2012).

Para o Ibram, o painel de visualização oferece recursos para uma gestão da documentação museológica mais eficaz, pois apresenta, de forma unificada e interativa,

uma visão macro de todos os seus museus, por meio de gráficos, métricas, tabelas e mapas.

Além disso, dados integrados podem demonstrar inconsistências sintáticas ou semânticas, possibilitando aos museus melhorarem, na base, suas informações a partir de boas práticas de catalogação (GILLILAND, 2016), incluindo criação e uso de metadados, de vocabulários controlados (ou linguagens documentárias) e de regras de catalogação, com as quais se determinam como elaborar a descrição de um recurso de informação e os pontos de acesso, tornando-se práticas essenciais na padronização e na descrição, viabilizando, portanto, interoperabilidade e agregação de acervos digitais. Nesse viés, instituições de memória cumprem a função de fornecer uma melhor experiência de busca para o usuário, potencializando um caráter formativo de sua equipe interna, tendo em vista que os museus são encorajados a disponibilizar conteúdo de qualidade na rede.

A partir da comparação de diferentes instituições, torna-se possível encontrar padrões e conexões que interessam (FERREIRA, 2012). Os indicadores podem garantir um melhor planejamento estratégico, maior produtividade do trabalho, agilidade no acesso aos dados, melhorando a eficiência da gestão, gerando efeitos nas dimensões administrativas, culturais e educacionais.

O painel também pode, a critério do Ibram, ficar disponível ao público em geral, permitindo que pesquisadores, por exemplo, possam enriquecer sua fonte de dados para pesquisa; que os usuários possam ter acesso aos links dos objetos digitais por meio de simples pesquisas; que os museus possam monitorar o seu acervo em comparação aos demais. As possibilidades de análises e públicos-alvo são inúmeras, pelo fato de o painel abrir um leque de novas possibilidades, até então, inimagináveis se considerarmos cada museu de forma independente.

Desta forma, o presente artigo tem como objetivo descrever o processo de criação do Painel de visualização analítica dos acervos digitais integrados do Ibram, utilizando o Tainacan e a pilha Elastic Stack, com ênfase no Kibana, assim como elucidar e apresentar o protótipo desenvolvido.

2 METODOLOGIA

Em relação aos procedimentos metodológicos, a presente pesquisa foi classificada como sendo de natureza qualitativa, de cunho exploratório e descritivo envolvendo o processo de agregação dos acervos museológicos digitais do Ibram e o desenvolvimento de um painel de visualização analítica de acervos integrados, configurando como modalidade de estudo de caso. Como procedimento técnico, utilizou-se de pesquisa bibliográfica no intuito de fundamentar conceitos e fornecer sustentabilidade teórica ao estudo.

O Painel de Visualização Analítica contemplou 19 museus, totalizando 16.057 itens, listados no Quadro 1.

Quadro 1: Relação de Museus, suas coleções, links e quantidade de itens

N.	Museu	Link das coleções	N. de itens
1	Museu Casa da Hera	http://museucasadahera.acervos.museus.gov.br	1.050
2	Museu Casa da Princesa	http://museusibramgoias.acervos.museus.gov.br/museu-casa-da-princesa	996
3	Museu Casa de Benjamin Constant	http://museucasabenjaminconstant.acervos.museus.gov.br/acervo-museologico	983
4	Museu Casa Histórica de Alcântara	http://museucasahistoricadealcantara.acervos.museus.gov.br/objetos	631
5	Museu da Abolição	https://museudaabolicao.acervos.museus.gov.br/acervo_museologico/	301
6	Museu da Inconfidência	http://museudainconfidencia.acervos.museus.gov.br/acervo-museologico	4.624
7	Museu das Bandeiras	http://museusibramgoias.acervos.museus.gov.br/museu-das-bandeiras	401
8	Museu das Missões	http://museudasmissoes.acervos.museus.gov.br/acervo-museologico	90
9	Museu de Arqueologia de Itaipu	http://museudearqueologiadeitaipu.museus.gov.br/museu-itaipu	1.040
10	Museu de Arte Religiosa e Tradicional	https://museudeartereligiosaetradicional.acervos.museus.gov.br/acervo-museologico	132
11	Museu de Arte Sacra da Boa Morte	http://museusibramgoias.acervos.museus.gov.br/museu-casa-da-boa-morte	783
12	Museu do Diamante	https://museudodiamante.acervos.museus.gov.br/acervo-museologico	895
13	Museu do Ouro	https://museudoouro.acervos.museus.gov.br/acervo	101
14	Museu Histórico Nacional	http://mhn.acervos.museus.gov.br/reserva-tecnica	847
15	Museu Regional Casa dos Ottoni	https://museuregionalcasadosottoni.acervos.museus.gov.br/acervo-museologico	463

16	Museu Regional do Caeté	https://museuregionaldecaete.acervos.museus.gov.br/acervo/	243
17	Museu Regional São João Del Rey	http://museuregionaldesaojoaodelrei.acervos.museus.gov.br/acervo_museologico	328
18	Museu Victor Meirelles	http://museuvictormeirelles.acervos.museus.gov.br/mvm-acervo	237
19	Museu Villa Lobos	http://museuvillalobos.acervos.museus.gov.br/fotografias	1.812
		TOTAL	**16.057**

Fonte: Elaborado pela autoria (2022)

De modo a integrar os acervos nos museus de origem e criar o painel de visualização foram utilizados o *software* Tainacan, versão 0.15, e a pilha do Elastic Stack, versão 7.9, ambas, versões mínimas para efetivação da integração. No caso do Tainacan, necessária para instalação do plugin de mapeamento de dados, e no Elastic Stack, mínima para que as configurações para controle da paginação feita pelo Tainacan tenham efeito. Todo código está disponível no GitHub, repositório tainacan-elk[4].

O Elastic Stack é um conjunto de quatro ferramentas, gratuitas e de código aberto, que possibilitam a obtenção de dados de diversas fontes e em qualquer formato, com vistas a buscar, a analisar e a visualizar dados em tempo real, a saber: Beats, Logstash, Elasticsearch e Kibana (ELASTIC, 2022b).

O Beats é uma plataforma para agentes de dados de finalidade única, que enviam dados de centenas ou milhares de computadores e sistemas para o Logstash ou o Elasticsearch (ELASTIC, 2022a). Nesta pesquisa, utilizamos o agente FileBeat. O Logstash é um pipeline de processamento de dados do lado do servidor que faz a ingestão de dados de inúmeras fontes, transforma-os e envia-os para o seu local de armazenamento (ELASTIC, 2022e). O Elasticsearch é um mecanismo de análise de dados e busca distribuído, que armazena os dados centralmente para proporcionar busca rápida (ELASTIC, 2022c). Por fim, o Kibana é uma interface de usuário para visualização dos dados do Elasticsearch e navegação no Elastic Stack (ELASTIC, 2022d).

[4] GITHUB. Tainacan - tainacan-elk. 2021. Disponível em: https://github.com/tainacan/tainacan-elk. Acesso em 14 mar. 2022

3 AGREGAÇÃO DOS DADOS DOS MUSEUS DO IBRAM

Para a criação do painel de visualização torna-se necessário, de forma geral, realizar a coleta, o processamento e o armazenamento dos dados, integrando os 19 museus. Estas etapas são descritas a seguir, subdivididas pelas tecnologias Tainacan e Elastic Stack.

3.1 Tainacan

Agregar acervos digitais nos remete a questões relacionadas à interoperabilidade, conforme supracitado por National Information Standards Organization (2004); Santarém Segundo, Silva e Martins (2019) e Innocenti, Vullo e Ross (2010).

Dando prosseguimento ao assunto, Arms (2002) citado por Fusco (2010), explicam que 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.

Neste ponto, 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 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*.

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.

No caso do Ibram, em 2014, a Resolução Normativa n. 2 estabeleceu os elementos de descrição das informações sobre o acervo museológico, bibliográfico e arquivístico, intitulado Inventário Nacional dos Bens Culturais Musealizados (INBCM) (BRASIL, 2021). Dessa forma, os elementos de metadados constituintes dos esquemas

das bases de dados dos museus envolvidos, sempre que possível, foram mapeados para o INBCM.

O mapeamento foi realizado diretamente no Tainacan por meio de um *plugin* extra (GITHUB, 2021), desenvolvido especificamente para esta pesquisa, que adiciona o mapeador do INBCM. O *plugin*, que requer a versão 0.15 do Tainacan ou superior, precisa ser instalado separadamente, pois não faz parte da versão principal do *software*, e está disponível no repositório tainacan-mappers-ibram-inbcm[5] do GitHub.

Para demonstração de como se deu o mapeamento, utilizaremos como exemplo o Museu Villa-Lobos, que possui 27 metadados, porém, apenas alguns são demonstrados no Quadro 2.

Quadro 2: Exemplos de mapeamento do Museu Villa Lobos para o INBCM

INBCM	**Museu Villa Lobos**	
Metadados	**Metadados**	**Status**
Número de registro	Número de registro	Público
Outros números	Outros números	Privado
Dimensões - altura	Sem representação	Público
Sem representação	Data padronizada	Privado
Sem representação	Localização	Público

Legenda:
- Mapeado - coletado
- Mapeado - não coletado
- Não mapeado - não coletado

Fonte: Elaborado pela autoria (2021).

No Tainacan, os usuários são livres para modelar quaisquer metadados que desejar e determinar se eles são públicos ou privados, ou seja, respectivamente, se serão disponibilizados ao público ou ficarão restritos à administração do museu. Assim, conforme o Quadro 2, apenas os metadados públicos e mapeados serão coletados e serão apresentados no painel de visualização.

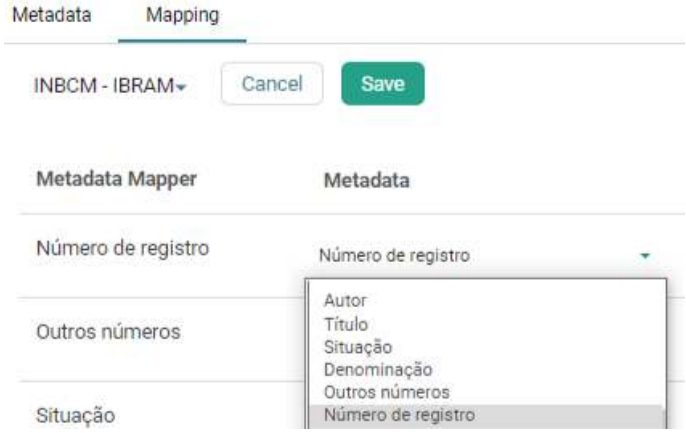
Além do INBCM, são incluídos três novos metadados no processo de agregação: o nome do museu de origem dos dados; um código único gerado pelas ferramentas do Elastic Stack e o link do item, que direciona ao objeto digital no museu de origem.

A Figura 1 mostra parte da interface de mapeamento do Tainacan, que, de forma simplificada, permite que os próprios museus executem o mapeamento.

[5] GITHUB. Tainacan. tainacan-mappers-ibram-inbcm. Disponível em: https://github.com/tainacan/tainacan-mappers-ibram-inbcm. Acesso em 14 mar. 2022.

Após o mapeamento, o Tainacan passa a disponibilizar a *Application Programming Interface* - API no formato *Java Script Object Notation* - JSON[6], paginados a cada 96 itens, que poderão ser coletados pelo *Elastic Stack*. A API é coletada pela pilha Elastic Stack, dando início ao processo de agregação dos dados.

Figura 1: Interface de mapeamento de dados do Tainacan



Fonte: elaborado pela autoria (2021).

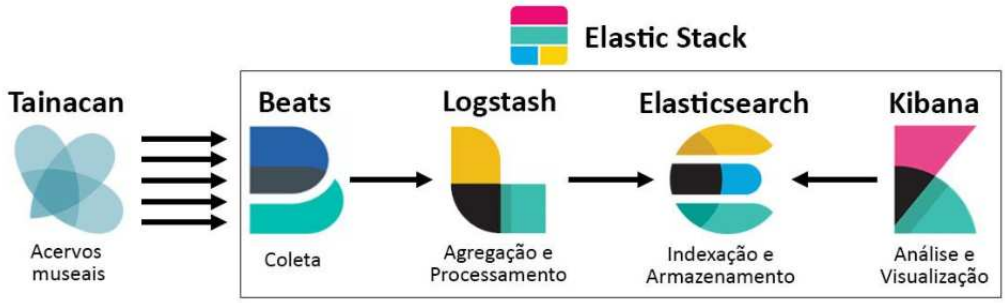
Para acessar, basta abrir a página web do Acervo e seguir os passos: Ver Como -> JSON Simples -> JSON simples: inbcm-ibram mapper. Como exemplo, o link da Página 1 do Museu Villa-Lobos[7].

3.2 Elastic Stack

Conforme supracitado, cada instalação do Tainacan disponibiliza a API com um retorno em JSON, e estes são coletados pelo Filebeat, dando início a agregação dos dados. A Figura 2 mostra a sequência.

[6] JSON.ORG. Introducing JSON. Disponível em: https://www.json.org/json-en.html. Acesso em 14 mar. 2022.
[7] JSON. Museu Villa-Lobos. Disponível em: https://museuvillalobos.acervos.museus.gov.br/wp-json/tainacan/v2/collection/1570/items/?perpage=96&order=DESC&orderby=date&exposer=json-flat&mapper=inbcm-ibram&paged=1. Acesso em 14 mar. 2022.

Figura 2: Etapas para integração das instalações do Tainacan com a Elastic Stack



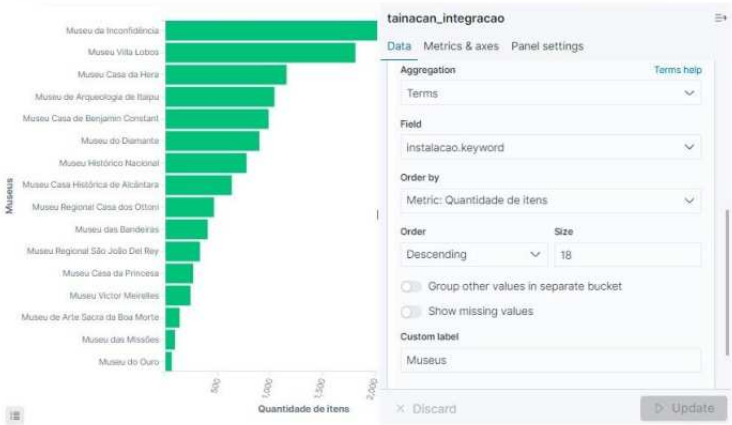
Fonte: elaborada pela autoria (2022).

O Beats (FileBeat) coleta os dados; o Logstash agrega, processa e realiza transformações importantes nos dados, como exemplo, retirar espaços em branco em excesso. Na sequência, tudo é armazenado no Elasticsearch, que abastece o Kibana com os dados que serão transformados em gráficos, tabelas, mapas, no painel de visualização.

4 PAINEL DE VISUALIZAÇÃO ANALÍTICA DOS ACERVOS DIGITAIS INTEGRADOS

A partir dos dados integrados, o Kibana possibilita a criação do painel e, para tal, se baseia na criação de elementos de visualização, tais como, gráficos, títulos, métricas, para compor o painel. Para exemplificar, apresentamos, na Figura 3, a interface de criação da visualização 'Horizontal Bar', ou seja, um gráfico de barras horizontais que demonstra a quantidade de itens por museu.

Figura 3: Interface do Kibana para criação de um gráfico com o número de itens dos museus

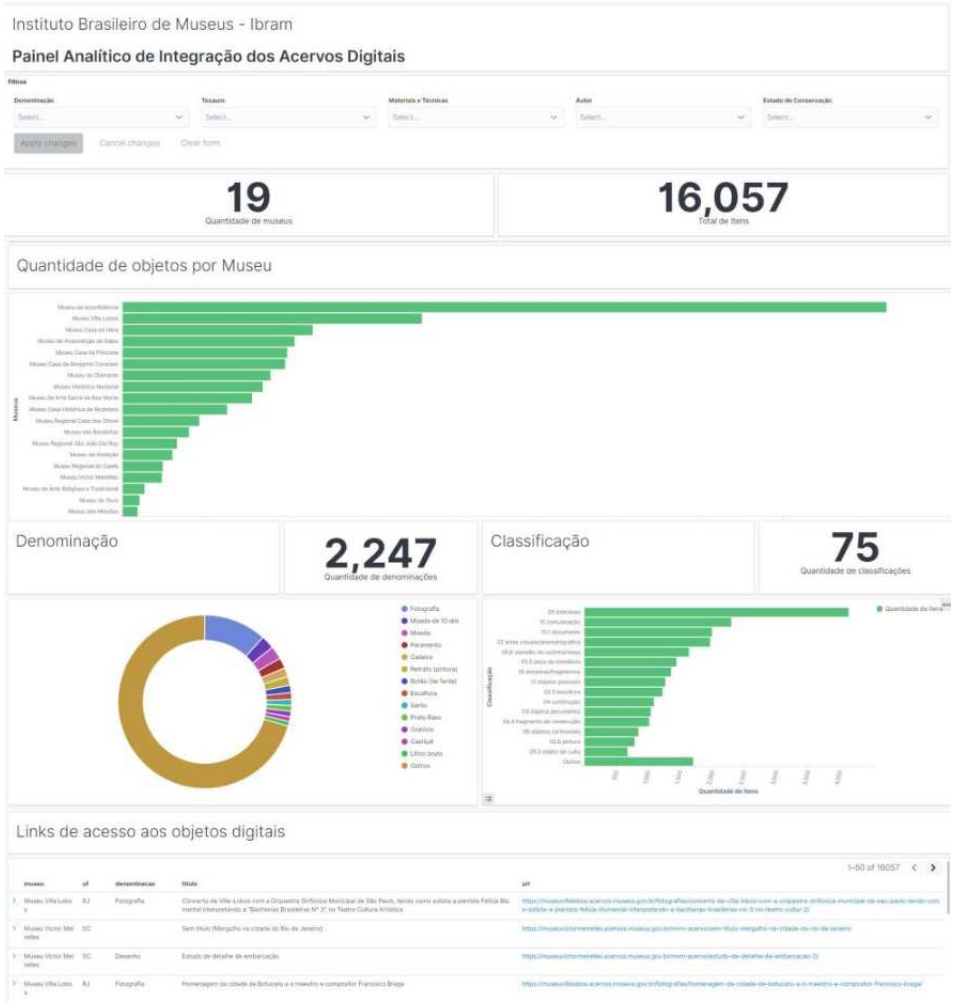


Fonte: elaborada pela autoria (2021).

Dessa forma, um painel é composto por um conjunto de visualizações. No caso do Ibram, são 30: nove títulos, cinco filtros, oito métricas numéricas, seis gráficos, uma tabela, um quadro de seleção, podendo, a qualquer momento, adicionar novas visualizações, a depender das necessidades da Instituição.

Em números são 19 museus, 16.057 itens, 2.247 denominações, 75 classificações, 755 autores, seis estados de conservação, 744 materiais e técnicas e 305 locais de produção. A Figura 4 apresenta parte do painel.

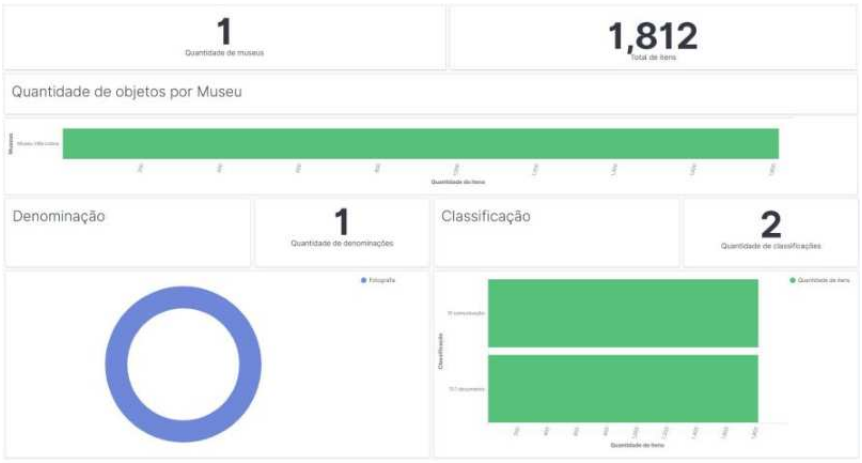
Figura 4: Painel de Visualização Analítica dos Acervos Digitais Integrados do Ibram



Fonte: elaborada pela autoria (2022).

Por fim, demonstramos algumas formas de realizar as buscas no painel de forma rápida e eficiente. A primeira, é possível interagir com os gráficos. Por exemplo, ao clicar em Museu Villa-Lobos, no gráfico de museus, todo painel se altera, mostrando apenas os números referentes a este museu, conforme demonstra a Figura 5.

Figura 5: Exemplo com o resultado da seleção do Museu Villa-Lobos no gráfico de museus



Fonte: elaborado pela autoria (2022).

Outra maneira são os filtros pré-configurados, que permitem a obtenção de resultados que combinem diferentes seleções. A Figura 6 apresenta o exemplo de busca por itens cujas denominações selecionadas são ‘oratório’ ou ’cadeira’, no metadado materiais e técnicas foi selecionada a ‘madeira’ e no metadado estado de conservação foi selecionado ‘bom’.

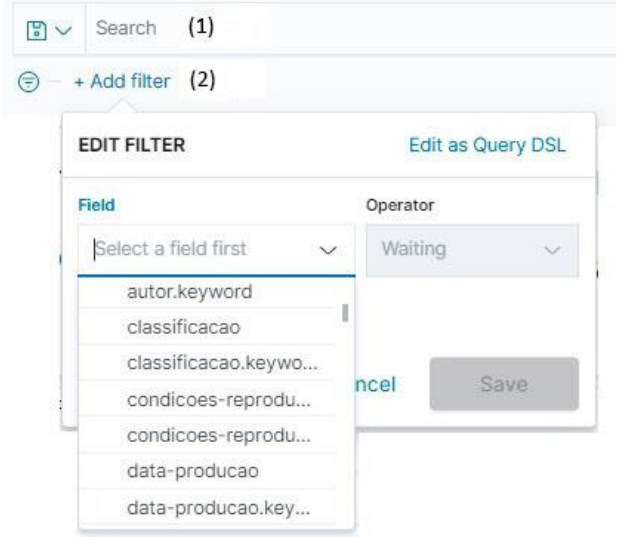
Figura 6: Exemplo de consulta com filtro pré-configurado



Fonte: elaborada pela autoria (2021).

O Kibana também permite realizar buscas que independem das visualizações pré-definidas, adicionando e combinando diferentes preferências (2), ou uma busca simples, por palavra-chave (1), conforme Figura 7.

Figura 7: Interface para uma consulta simples, por palavra-chave e com novos filtros



Fonte: elaborada pela autoria (2021).

O painel criado utilizando o Kibana abre diversas possibilidades, e permite aos gestores um controle de todos os acervos, em uma única interface. Por ser utilizado junto a Elastic Search, ao FileBeat e ao Logstash, as atualizações realizadas nos museus de origem serão automaticamente atualizadas. Além disso, permite a criação de diversos outros gráficos, tabelas, a partir das necessidades apresentadas pelo Ibram, tornado o painel ainda mais completo

Para visualizar o resultado final e realizar as consultas desejadas, disponibilizamos o acesso para o painel de visualização analítica dos acervos integrados dos museus vinculados ao Ibram, disponível no link: http://200.137.217.144:6601/goto/613663b72fc306a908e86a6233e5219c

5 CONSIDERAÇÕES FINAIS

Esta pesquisa resultou na criação do Painel de Visualização Analítica dos Acervos Integrados do Ibram, a partir da utilização de duas diferentes tecnologias, o *software* Tainacan e a pilha Elastic Stack. Os resultados obtidos são parte de uma pesquisa de

doutoramento de maior amplitude, visando a construção de um agregador de objetos culturais de museus brasileiros a ser denominado Brasiliana Museus.

Considera-se que essa etapa de criar repositórios digitais de dados da documentação museológica, tratar os dados, mapear para um modelo comum, coletar e gerar uma base integrada de dados para o cálculo de indicadores e a geração de painéis de visualização de dados foi concluída de forma satisfatória. Nesse sentido, tornou-se possível a geração de informações úteis que têm ajudado a compreender a documentação dos museus de uma forma ainda não trabalhada pelo Ibram.

A pesquisa demonstra a viabilidade das tecnologias para agregação de dados e, portanto, valida o modelo tecnológico e o fluxo de circulação da informação para a continuidade da pesquisa em busca da modelagem e da implementação da ferramenta de busca integrada dos acervos. Importante destacar a robustez do Elasticsearch, muito eficaz na busca de texto completo, com escalabilidade horizontal capaz de manipular penta bytes de dados, garantindo a agregação de milhares de itens museológicos.

Vale ressaltar que a combinação do Tainacan e da pilha do Elastic Stack se mostrou de média complexidade, se tornando uma opção viável de aprendizagem e customização para o uso e a análise de dados oriundos de repositórios digitais.

REFERÊNCIAS

AVAZPOUR, Iman; GRUNDY, John; ZHU, Liming. Engineering complex data integration, harmonization and visualization systems. **Journal of Industrial Information Integration**, [*s.l.*], v. 16, dec. 2019. Disponível em: https://doi.org/10.1016/j.jii.2019.08.001. Acesso em: 15 dez. 2022.

BRASIL. Instituto Brasileiro de Museus. Resolução Normativa IBRAM n. 6, de 31 de agosto de 2021. **Diário Oficial da União**, Brasília, 2021. Disponível em: https://www.in.gov.br/web/dou/-/resolucao-normativa-ibram-n-6-de-31-de-agosto-de-2021-342359740. Acesso em 14 mar. 2022.

BRASIL. Ministério do Turismo. **Sobre o órgão**. Brasília, 2022. Disponível em: https://www.gov.br/museus/pt-br/acesso-a-informacao/institucional/sobre-o-orgao. Acesso em 14 mar. 2022.

BRASIL. **Projeto Tainacan**. Belo Horizonte: Ministério do Turismo; Instituto Brasileiro de Museus (Ibram), 2021. Disponível em: https://www.gov.br/museus/pt-br/acesso-a-

informacao/acoes-e-programas/acervo-em-rede-e-projeto-tainacan. Acesso em 14 mar. 2022.

BRASILIANA FOTOGRÁFICA. **Documentos técnicos**. [*S.l.*], 2022c. Disponível em: https://brasilianafotografica.bn.gov.br/?page_id=5329. Acesso em 21 fev. 2022.

BRASILIANA FOTOGRÁFICA. **Objetivos**. [*S.l.*], 2022a. Disponível em: https://www.brasilianaiconografica.art.br/sobre-o-projeto. Acesso em 21 fev. 2022.

BRASILIANA FOTOGRÁFICA. **Quem somos.** [*S.l.*], 2022b. Disponível em: https://brasilianafotografica.bn.gov.br/?page_id=119. Acesso em 21 fev. 2022.

BRASILIANA ICONOGRÁFICA. **Sobre o Projeto**. [*S.l.*], 2017. Disponível em: https://www.brasilianaiconografica.art.br/sobre-o-projeto. Acesso em 21 fev. 2022.

CHAN, Lois Mai; ZENG, Marcia Lei. Metadata Interoperability and Standardization – A Study of Methodology Part I. Achieving Interoperability at the Schema Level. **D-Lib Magazine**, [*s.l.*], v. 12, n. 6, June 2006. ISSN 1082-9873. Disponível em: https://doi.org/10.1045/june2006-chan. Acesso em: 15 dez. 2022.

ELASTIC. **Beats**. Agentes de dados lightweight. [*S.l.*], 2022a. Disponível em: https://www.elastic.co/pt/beats/. Acesso em: 14 mar. 2022.

ELASTIC. **Elastic Stack**. [*S.l.*], 2022b. Disponível em: https://www.elastic.co/pt/elastic-stack. Acesso em: 14 mar. 2022.

ELASTIC. **Elasticsearch**. O coração do Elastic Stack gratuito e aberto. [*S.l.*], 2022c. Disponível em: https://www.elastic.co/pt/Elasticsearch/. Acesso em: 14 mar. 2022.

ELASTIC. **Kibana**. A sua janela para o Elastic Stack. [*S.l.*], 2022d. Disponível em: https://www.elastic.co/pt/kibana. Acesso em: 14 mar. 2022.

ELASTIC. **Logstash**. Centralize, transforme e oculte seus dados. [*S.l.*], 2022e. Disponível em: https://www.elastic.co/pt/logstash. Acesso em: 14 mar. 2022.

FERREIRA, Nuno Miguel Antunes. **Visualização de informação e visualização analítica:** mapa de visualização gráfica da informação agregada do país, um sistema de apoio à decisão. 2012. Dissertação (Mestrado em Sistemas Integrados de Apoio à Decisão) - Instituto Universitário de Lisboa. Escolas de Tecnologias e Arquitetura. Departamento de Ciências e Tecnologias da Informação. Portugal. Disponível em: https://repositorio.iscte-iul.pt/handle/10071/5906. Acesso em: 14 mar. 2022.

FREIRE, Nuno, *et al*. **Web Technologies**: A Survey of Their Applicability to Metadata Aggregation in Cultural Heritage. Expanding Perspectives on Open Science: Communities, Cultures and Diversity in Concepts and Practices, 2017. p. 235-244. Disponível: https://doi.org/10.3233/978-1-61499-769-6-235. Acesso em: 16 dez. 2022.

FUSCO, Elvis. **Modelos conceituais de dados como parte do processo de catalogação**: perspectiva do uso dos FRBR no desenvolvimento de catálogos bibliográficos digitais. Tese de Doutorado. Marília: Unesp, 2010. Disponível em: http://www.marilia.unesp.br/Home/Pos-Graduacao/CienciadaInformacao/Dissertacoes/fusco_e_do_mar.pdf. Acesso em: 04 mar. 2022.

GÄNGLER, Thomas. **Semantic Federation of Musical and Music-Related Information for Establishing a Personal Music Knowledge Base**. 2011. Tese (Doutorado em Ciência da Computação) - Department of Computer Science, Institute for Systems Architecture, Chair of Computer Networks, Technische Universität Dresden, Dresden – Alemanha, 2011. Disponível em: https://nbn-resolving.org/urn:nbn:de:bsz:14-qucosa-72434. Acesso em 03 mar. 2022.

GILLILAND, Anne J. Setting the Stage. *In*: BACA, M. (Ed.). **Introduction to metadata**. 3. ed. Los Angeles: Getty Research Institute, 2016. Disponível em: https://www.getty.edu/publications/intrometadata/setting-the-stage/. Acesso em 14 mar. 2022.

GITHUB. **Tainacan**. tainacan-mappers-ibram-inbcm. [*S.l.*], 2021. Disponível em: https://github.com/tainacan/tainacan-mappers-ibram-inbcm. Acesso em 14 mar. 2022.

GOVERNO DO RIO DE JANEIRO. Secretaria de Estado da Cultura. **Museus**. Rede Web de Museus. Rio de Janeiro, 2022b. Disponível em: http://www.museusdoestado.rj.gov.br/rede-web-de-museus/. Acesso em 21 fev. de 2022.

INSTITUTO BRASILEIRO DE MUSEUS. **Sobre o órgão**. Belo Horizonte, 2022. Disponível em: https://www.gov.br/museus/pt-br/acesso-a-informacao/institucional/sobre-o-orgao. Acesso em: 04 mar. 2022.

INNOCENTI, Perla; VULLO, Giuseppina; ROSS, Seamus. Towards a Digital Library Policy and Quality Interoperability Framework: The DL.org Project. **New Review of Information Networking**, [*s.l.*], v. 15, n. 1, p. 29-53, 2010. Disponível em: https://doi.org/10.1080/13614571003751071. Acesso em: 14 dez. 2022.

KIRSTEN, Toralf, *et al*. Metadata **Management for Data Integration in Medical Sciences**. Datenbanksysteme für Business, Technologies and Web (BTW 2017), 2017. Disponível em: https://dl.gi.de/handle/20.500.12116/627. Acesso em 04 mar. 2022.

MARCONDES, Carlos Henrique. Interoperabilidade entre acervos digitais de arquivos, bibliotecas e museus: potencialidades das tecnologias de dados abertos interligados. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 21, n. 2, p. 61-83, jun. 2016. Disponível em: http://portaldeperiodicos.eci.ufmg.br/index.php/pci/article/view/2735. Acesso em 03 mar. 2022.

MARTINS, Dalton Lopes; CARVALHO JÚNIOR, José Murilo Costa; GERMANI, Leonardo. Projeto Tainacan: experimentos, aprendizados e descobertas da cultura digital no universo dos acervos das instituições memoriais. *In*: COMITÊ GESTOR DA INTERNET. **TIC cultura 2018** - Pesquisa Sobre o Uso das Tecnologias de Informação e Comunicação nos Equipamentos Culturais Brasileiros. Núcleo de Informação e Coordenação do Ponto BR. 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 14 mar. 2022.

MARTINS, Dalton Lopes; SILVA, Marcel Ferrante; CARMO, Danielle. Acervos em rede: perspectivas para as instituições culturais em tempos de cultura digital. **Em Questão**, Porto Alegre, v. 24, n. 1, p. 194-216, 2018. Disponível em: https://doi.org/10.19132/1808-5245241.194-216. Acesso em: 15 dez. 2022.

MARTINS, Luciana Conrado; MARTINS, Dalton Lopes. Experimentações sociotécnicas para organização e difusão de coleções digitais universitárias: o caso do projeto Tainacan. **Revista CPC**, [*s.l.*], v. 15, n. 30, p. 34-61, 2020. Disponível em: https://doi.org/10.11606/issn.1980-4466.v15i30espp34-61. Acesso em: 15 dez. 2022.

MILES, Alistair; ZHAO, Jun; KLYNE, Graham, WHITE-COOPER, Helen; SHOTTON, David. OpenFlyData: an exemplar data web integrating gene expression data on the fruit fly Drosophila melanogaster. **Journal of biomedical informatics**, [*s.l.*], v. 43, n. 5, p. 752-761, 2010. Disponível em: https://doi.org/10.1016/j.jbi.2010.04.004. Acesso em: 15 dez. 2022.

MUCHACHO, Rute. **O Museu Virtual: as novas tecnologias e a reinvenção do espaço museológico**. Biblioteca online de ciências da comunicação, 2005. Disponível em: http://www.bocc.ubi.pt/pag/muchacho-rute-museu-virtual-novas-tecnologias-reinvencao-espaco-museologico.pdf. Acesso em: 03 de mar. 2022

NATIONAL INFORMATION STANDARDS ORGANIZATION. **Understanding Metadata**: What is Metadata, and What is it For? A Primer. Bethesda, MD: NISO Press. 2004. Disponível em: http://www.niso.org/standards/resources/UnderstandingMetadata.pdf. Acesso em 03 mar. 2022.

PIERRE, Margaret St; LAPLANT, William P. **Issues in crosswalking content metadata standards**. Baltimore, Maryland, USA: NISO, 1998. Disponível em: https://groups.niso.org/publications/white_papers/crosswalk/. Acesso em: 03 mar. 2022

SANTAREM SEGUNDO, José Eduardo; SILVA, Marcel Ferrante; MARTINS, Dalton Lopes. Revisitando a interoperabilidade no contexto dos acervos digitais**. Informação & Sociedade**: Estudos, João Pessoa, v. 29, n. 2, 2019. Disponível em: https://periodicos.ufpb.br/ojs2/index.php/ies/article/view/38107. Acesso em: 25 jan. 2022.

SARIKAYA, Alper, *et al*. What Do We Talk About When We Talk About Dashboards? **IEEE Transactions on Visualization and Computer Graphics**, [*s.l.*], v. 25, n. 1, pp. 682-692, 2019. Disponível em: https://doi.org/10.1109/TVCG.2018.2864903. Acesso em: 15 dez. 2022.

SIQUEIRA, Joyce; MARTINS, Dalton Lopes. Recuperação de informação: descoberta e análise de workflows para agregação de dados do patrimônio cultural. **Ciência da Informação**, Belo Horizonte, v. 49, n. 3, 2020. Disponível em: http://revista.ibict.br/ciinf/article/view/5399. Acesso em: 20 jan. 2022

SIQUEIRA, Joyce; MARTINS, Dalton Lopes. Workflow de agregação de dados: Processos para criação de uma interface de busca integrada do patrimônio cultural, 2019a. Workshop de informação, dados e tecnologia. *In*: WORKSHOP DE INFORMAÇÃO, DADOS E TECNOLOGIA - WIDAT, 2019, Brasília. **Anais** [...]. Brasília: Universidade de Brasília, Faculdade de Ciência da Informação, 2019. Disponível em: http://widat2019.fci.unb.br/index.php/anais-widat-2019. Acesso em: 03 mar. 2022.

SIQUEIRA, Joyce; MARTINS, Dalton Lopes. Workflow models for aggregating cultural heritage data on the web: a systematic literature review, 2021b. **Journal of the Association for Information Science and Technology**, [*s.l.*], v. 73, n. 2, p. 204-224, 2021. Disponível em: https://doi.org/10.1002/asi.24498. Acesso em: 14 dez. 2022.

TAINACAN. **Casos de Uso**. 2022. Disponível em: https://tainacan.org/casos-de-uso/. Acesso em: 14 mar. 2022.

ZIÉBELIN, Danielle, *et al*. Heterogeneous Data Integration Using Web of Data Technologies, 2017. *In*: WEB AND WIRELESS GEOGRAPHICAL INFORMATION SYSTEMS - W2GIS 2017. 2017, [*s.l.*]. **Anais** [...]. [*s.l.*]: Springer, v. 10181, 2017. Disponível em: https://doi.org/10.1007/978-3-319-55998-8_3. Acesso em: 14 dez. 2022.