

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(Ibram)

Danielle do Carmo

Doutoranda em Ciência da Informação pela Universidade de Brasília (UnB) - Brasília, DF - Brasil. Mestre em Memória Social e Patrimônio Cultural pela Universidade Federal de Pelotas (UFPEL) – RS - Brasil.

<http://lattes.cnpq.br/6139212172511823>

E-mail:docarmo.danielle@gmail.com

Dalton Lopes Martins

Pós-Doutorado pela Universidade de São Paulo (USP) – SP - Brasil. Doutor em Ciência da Informação pela Universidade de São Paulo (USP) - São Paulo, SP - Brasil. Professor da Universidade de Brasília (UnB) - Brasília, DF - Brasil.

<http://lattes.cnpq.br/3774617443225038>

E-mail:dmartins@gmail.com

Submetido em: 23/09/2020. Aprovado em: 25/11/2020. Publicado em: 28/07/2021.

RESUMO

O artigo apresenta um estudo webométrico acerca de mídias referentes a coleções de acervos de museus brasileiros no repositório de mídias Wikimedia Commons, mais especificamente sobre mídias de coleções de nove museus geridos pelo Instituto Brasileiro de Museus (Ibram). Utilizando as ferramentas GLAMorous e GLAMorous 2, foi possível obter dados relativos ao reúso de mídias da Wikimedia Commons em outras plataformas wiki da Fundação Wikimedia, como os projetos Wikipédia de diferentes idiomas, a Wikidata e outros. Também foi possível obter números relativos à quantidade de visualizações desses acervos no ano de 2019. Dessa forma a investigação realizada, assim como os dados obtidos, podem ajudar instituições culturais guardiãs de acervos, como os museus, a entender o que acontece com seus acervos digitais uma vez que são disponibilizados nas plataformas wiki, fornecendo dados importantes para a gestão de suas coleções, principalmente em relação ao acesso e reapropriação desses acervos por parte dos usuários.

Palavras-chave: Wikimedia Commons. Acervos culturais digitais. Wikidata. Instituições culturais. Museus. GLAM.

Brazilian Cultural Collections in the Wikimedia Commons Repository: a study on the reuse and visualization of media related to museum collections of the Brazilian Institute of M

useums (Ibran)

ABSTRACT

This article presents a webometric study about media referring to collections of Brazilian museums in the Wikimedia Commons media repository, more specifically about media from collections of nine museums managed by the Brazilian Museum Institute (Ibran). Using the GLAMorous and GLAMorous 2 tools, it was possible to obtain data regarding the media reuse of Wikimedia Commons in other Wikimedia Foundation wiki platforms, such as Wikipedia projects in different languages, Wikidata and others. It was also possible to obtain numbers related to the views of these collections in 2019. In this way, the research carried out, as well as the data obtained, ways to help cultural institutions guarding collections, such as museums, to understand what happens with their digital collections. Since they are made available on wiki platforms, providing important data for the management of their collections, mainly in relation to access and reappropriation of these collections by users.

Keywords: *Wikimedia Commons. Digital cultural collections. Wikidata. Museums. Cultural institutions. GLAM.*

Colecciones culturales brasileñas en el repositorio de Wikimedia Commons: un estudio sobre la reutilización y visualización de medios relacionados con colecciones de museos del Instituto Brasileño de Museos(Ibran)

RESUMEN

Este artículo presenta un estudio webométrico sobre los medios que se refieren a colecciones de museos brasileños en el repositorio de medios de Wikimedia Commons, más específicamente sobre medios de colecciones de nueve museos administrados por el Instituto Brasileño de Museos (Ibran). Utilizando las herramientas GLAMorous y GLAMorous 2, fue posible obtener datos sobre la reutilización de medios de Wikimedia Commons en otras plataformas de wikis de la Fundación Wikimedia, como proyectos de Wikipedia en diferentes idiomas, Wikidata y otros. También fue posible obtener números relacionados con el número de vistas de estas colecciones en 2019. De esta manera, la investigación llevada a cabo, así como los datos obtenidos, pueden ayudar a las instituciones culturales que custodian colecciones, como los museos, a comprender lo que sucede con sus colecciones digitales. ya que están disponibles en plataformas wiki, proporcionando datos importantes para la gestión de sus colecciones, principalmente en relación con el acceso y la reapropiación de estas colecciones por parte de los usuarios.

Palabras clave: *Wikimedia Commons. Colecciones culturales digitales. Wikidata. Instituciones culturales. Museos. GLAM.*

INTRODUÇÃO

Com o surgimento das novas tecnologias de informação e comunicação, mais especificamente com o advento da internet, pôde-se observar a emergência de novos espaços sociais baseados no digital e em novos tipos de fluxos informacionais. Esses espaços, ocupados com objetivos e propósitos diversos, oferecem aos seus usuários meios de consumo e compartilhamento de informações, mecanismos de interação com as informações e dispositivos que promovem a conexão e comunicação direta entre usuários. Como exemplo desses meios sociais digitais que emergiram principalmente com o surgimento da Web 2.0, podemos citar as plataformas baseadas em conteúdos gerados pelo usuário, como a rede social Facebook, a rede de micro blogs Twitter, o streaming de vídeo YouTube e a enciclopédia livre Wikipédia. Esses espaços sociotécnicos, além de serem entendidos como meios potenciais de socialização da informação, também podem ser entendidos como fontes de informação que permitem a coleta de dados sobre a circulação e a apropriação de objetos digitais, que nos fornece informações sobre como são disponibilizados, contextualizados e descritos.

Nesse sentido, a webometria pode ser uma importante aliada quando utilizada em estudos que pretendem investigar fenômenos baseados na web. Os autores Björneborn e Ingwersen (2004) definiram a webometria como o “estudo de fenômenos da web baseados em técnicas quantitativas e recorrendo a métodos infométricos” (2004). Para Thelwall (2009, p. 1) essa definição foi importante por atribuir à webometria características de método infométrico, pois dessa forma os autores a posicionaram como um campo da ciência da informação. Entretanto, o autor propõe uma definição mais abrangente ao dizer que a webometria é “o estudo dos conteúdos da web com métodos essencialmente quantitativos que não são específicos para um campo de estudo” (THELWALL, 2009, p. 6).

Vanti (2002, p. 161) posiciona a webometria, a bibliometria e a cienciometria como subcampos da infometria e diz que todas “têm funções semelhantes, mas ao mesmo tempo cada uma delas propõe medir a difusão do conhecimento científico e o fluxo da informação sob enfoques diversos” (VANTI, 2002). Entre os estudos webométricos estariam incluídos aqueles que adotam métodos de rastreamento das ações dos usuários on-line que, por meio de ferramentas de análise, podem captar informações que nos permitem medir aspectos do uso de recursos disponíveis na web (THELWALL, 2009, p. 89).

Do mesmo modo que os estudos bibliométricos permitem melhor administração das coleções no contexto das bibliotecas, os estudos webométricos podem dar subsídios que auxiliem as instituições culturais como os museus a administrar seus recursos, entre eles objetos digitais de acervos culturais, e entender como se dá seu uso pelos usuários que a eles têm acesso. Logo, a presente investigação explora algumas ferramentas de coleta de dados e apresenta possibilidades de análise e mensuração de aspectos relacionados ao acesso e à reutilização de acervos brasileiros que estão disponíveis na web por meio do repositório de mídias Wikimedia Commons.

O repositório de mídias on-line Wikimedia Commons armazena e disponibiliza, gratuitamente, diversos tipos de arquivos de mídias que são fornecidos de forma coletiva e colaborativa por meio da ação de usuários voluntários. É um projeto da organização sem fins lucrativos Fundação Wikimedia, que também é a responsável pela manutenção de outros projetos baseados em conteúdo gerado por usuários como a Wikipédia, o Wikitionary, o Wikibooks, o Wikisource, o Wikinews, o Wikiversite, o Wikiquote, o Wikidata e outros.

O Wikimedia Commons, lançado no início de setembro de 2004, disponibiliza atualmente mais de 60 milhões de arquivos de mídias. Essas mídias são disponibilizadas sob licenças individuais que permitem a cópia, a reutilização e a modificação, de acordo com termos especificados.

Os conteúdos disponibilizados por meio do repositório Wikimedia Commons estão sob a licença Creative Commons Attribution / Share-Alike (WIKIMEDIA COMMONS, 2020a). Segundo dados fornecidos pelo projeto, em abril de 2020, a plataforma registra um total de 43.020 usuários ativos e 168 *bots*, *programas que podem editar ou submeter conteúdo para a plataforma de modo automático*, que auxiliam na edição do conteúdo (WIKIMEDIA COMMONS, 2020b).

Ao explorar o repositório Wikimedia Commons, é possível observar que estão disponíveis diferentes tipos mídias de diversas temáticas. Entre as mídias é possível identificar conteúdos referentes a itens de acervos oriundos de coleções de diversas instituições culturais como museus, arquivos e galerias localizadas em diversas partes do mundo. Essas mídias podem ter sido disponibilizadas tanto por usuários que realizaram o upload desse material de modo espontâneo, ou em alguns casos por meio de parcerias estabelecidas entre usuários e instituições culturais. Segundo os autores Stinson, Fauconnier e Wyatt (2018, p. 17), há um histórico de colaboração entre a comunidade Wikimedia e as instituições culturais, reunidas sob o termo guarda-chuva GLAM¹. Essas instituições identificaram nos projetos da Fundação Wikimedia, em especial no Wikimedia Commons, na enciclopédia on-line Wikipédia e no repositório de dados estruturados Wikidata, meios potenciais de difusão de seus acervos na internet. Para se ter uma noção dos efeitos dessa prática, Zeinstra (2013) aponta que, no ano 2013, o conteúdo GLAM reunia cerca de mais de dois milhões de objetos digitais no Wikimedia Commons, o que corresponderia ao total de 13,14% do conteúdo da plataforma.

Em outras palavras, um em cada oito arquivos na Wikimedia Commons é disponibilizado através de alguma colaboração com GLAM's. Tanto a GLAM coloca suas coleções em domínio público ou abre a licença de suas coleções on-line e voluntários realizam upload, ou alguma colaboração ativa entre GLAM's e a comunidade Wikimedia como Wiki Loves Monuments cria conteúdo GLAM.

Dessa forma, voluntários colocam esses objetos de mídia em artigos da Wikipédia criando um incrível aumento em sua visibilidade. Instituições contribuintes enxergam o potencial da Wikimedia como um canal de distribuição (ZEINSTRAS, 2013, p. 5, tradução nossa).

As autoras Villaespesa e Navarete (2019) chamam a atenção para a utilização de mecanismos de busca como o Google e o Google Imagens, ou da busca de voz de assistentes virtuais como a Siri e a Alexa é possível identificar uma exposição privilegiada de conteúdos da Wikimedia Commons, Wikidata e Wikipédia. Essa exposição privilegiada na ordem de apresentação de resultados se daria devido “aos ambientes estruturados fornecidos pelas plataformas Wiki que influenciam fortemente os resultados das pesquisas” (VILLAESPESA; NAVARRETE, 2019).

Além de se apresentar como potenciais aliados na difusão das informações referentes aos acervos das instituições culturais, os projetos da Fundação Wikimedia constituem-se como espaços sociotécnicos que promovem práticas de curadoria coletiva e que podem nos fornecer dados relevantes sobre aspectos relacionados ao acesso e à reutilização das coleções de acervos culturais compartilhados nesses ambientes. Nesse sentido a Fundação Wikimedia, por meio de uma página na Wikimedia Outreach sobre os projetos com GLAM's, indica uma série de ferramentas desenvolvidas com o objetivo de captar e os expor dados sobre as mídias digitais de coleções culturais em ambientes como a Wikimedia Commons (WIKIMEDIA OUTREACH, 2020).

O objetivo da presente pesquisa foi explorar como esses dados podem ser extraídos e analisados, evidenciando seu potencial como fonte de informação sistematizada, assim como ampliar a própria sensibilização da área da ciência da informação para os fenômenos culturais que se dão nesses ambientes e o quanto eles podem indicar novas práticas sociais de gestão e curadoria da informação que devem ser levadas em consideração em pesquisas na área.

¹ Acrônimo em inglês - Galleries, Library, Archives and Museums - usado para se referir a instituições como galerias, bibliotecas, arquivos, museus e outras instituições do patrimônio cultural.

O foco da pesquisa é investigar a presença de acervos culturais na Wikimedia Commons e entender como se configura essa presença a partir de diferentes tipos de indicadores e informações disponibilizadas pela plataforma. A amostra utilizada foi de mídias de coleções de museus que estão sob a gerência do Instituto Brasileiro de Museus (Ibram).

METODOLOGIA

A pesquisa se apresenta como exploratória e descritiva, de natureza quantitativa, e para que fosse possível coletar os dados para análise, utilizaram-se duas ferramentas, a GLAMorous (MANSKE, 2020a) e a GLAMorous 2 (MANSKE, 2020b). Ambas as ferramentas foram desenvolvidas pelo wikimedista2 Magnus Manske e se encontram listadas em uma página de ferramentas voltadas para os GLAM's no Wikimedia Outreach (WIKIMEDIA OUTREACH, 2020). Com essas ferramentas foi possível obter dados de visualização dos arquivos das mídias do Wikimedia Commons com base nas categorias. Assim, depois de identificar os museus sob gestão do Ibram (2020), exploramos as categorias referentes às coleções destes museus na Wikimedia Commons. Uma vez identificadas as categorias, foi possível utilizá-las nos mecanismos de busca das ferramentas selecionadas e estabelecer os parâmetros. A ferramenta GLAMorous permite o rastreamento de uso de imagens que estão em uma categoria do Wikimedia Commons, e por meio dela foi possível obter dados relativos ao reuso dessas mídias em páginas de diversos projetos da Fundação Wikimedia. É importante destacar que a possibilidade de rastrear o uso de objetos digitais pode ser aplicada em pesquisas que tenham por objetivo compreender as formas de apropriação social e reuso da informação, sobretudo ressaltando o contexto em que elas são utilizadas; desse modo se torna possível analisar e evidenciar significados fundamentais para a compreensão dos fluxos informacionais constituídos nesses ambientes.

² Nome atribuído aos usuários e colaboradores dos diversos projetos da Fundação Wikimedia.

Desde o momento de publicação de um objeto digital por uma instituição, passando pelo momento em que esse objeto é escolhido para ilustrar um conceito em um verbete, até sua visualização por um usuário em um verbete específico na Wikipédia, diversas camadas podem compreendidas por meio da obtenção dessas informações registradas, podendo assim apoiar a compreensão do fenômeno social de uso da informação que ainda hoje carece de entendimento e fortalecimento metodológico. A ferramenta GLAMorous 2 complementa a GLAMorous, apresentando funcionamento semelhante, porém, além de dados de reuso, é possível obter dados de visualização dos arquivos. Com essas ferramentas foi possível realizar as análises webométricas que serão apresentados nas próximas seções do presente artigo.

ANÁLISE E DISCUSSÃO DOS RESULTADOS

A REUTILIZAÇÃO DE IMAGENS DE COLEÇÕES CULTURAIS NA WIKIMEDIA COMMONS EM OUTROS PROJETOS DA FUNDAÇÃO WIKIMEDIA

Ao identificar os 30 museus que estão sob a administração direta do Ibram, foi possível localizar categorias de mídias que correspondem às suas coleções no Wikimedia Commons por meio de buscas na plataforma³. Assim identificamos as categorias referentes às coleções de nove museus que estão elencadas no quadro 1.

Identificadas as categorias referentes a coleções de museus, foi possível utilizá-las na realização de buscas na ferramenta GLAMorous, e assim coletar estatísticas atuais de reuso de imagens da Wikimedia Commons em páginas da Wikipédia em português, em Wikipédias de outros idiomas e em páginas de outros projetos da Fundação Wikimedia, como o Wikibooks, Wikiquote e Wikidata. Além disso foi possível obter a quantidade de imagens encontradas em cada categoria, o total de imagens distintas utilizadas, o total de uso das imagens das categorias, assim como a porcentagem do total de imagens da categoria que estão sendo reutilizadas, como pode ser visto nas tabelas 1 e 2.

³ Busca por categorias: Disponível em: <<https://commons.wikimedia.org/wiki/Special:Categories>>. Acesso em abr. 2020.

Quadro 1 – Categorias referentes às coleções dos museus no Wikimedia Commons

#	Museu	Categoria
1	Museu da Inconfidência	Collections of the Museu da Inconfidência
2	Museu da República	Collections of the Museu da República
3	Museu Histórico Nacional	Collections of the Museu Histórico Nacional
4	Museu Imperial	Collections of the Museu Imperial
5	Museu Nacional de Belas Artes	Collections of the Museu Nacional de Belas Artes
6	Museu do Açude (equipamento dos Museus Castro Maya)	Collections of the Museu Castro Maya
7	Museu Casa de Benjamin Constant	Media contributed by the Museu Casa de Benjamin Constant
8	Museu Regional de São João del-Rei	Collections of the Museu Regional de São João del-Rei
9	Museu Victor Meirelles	Collections of the Museu Victor Meirelles

Fonte: Dados da pesquisa, 2020.

Com base nos dados obtidos, é possível observar que a coleção de museu que mais apresenta imagens reutilizadas na ilustração páginas da Wikipédia em português é a categoria referente ao Museu da República, que reúne 33 ocorrências de imagens presentes em artigos.

Essa categoria também é a que ilustra mais Wikipédias de outros idiomas, mostrando assim 75 ocorrências em páginas de Wikipédias em outros idiomas, uma ocorrência de imagem no Wikibooks e uma no Wikiquote.

Assim foi possível constatar que das nove imagens encontradas na categoria correspondente à coleção do Museu da República, 4 estão sendo reutilizadas em 110 páginas de projetos da Fundação Wikimedia, o que corresponde ao total de uso de 44,4% das imagens disponíveis na categoria.

A categoria referente à coleção do Museu Histórico Nacional revelou o maior número de arquivos, com 471 imagens encontradas na categoria. Essas imagens foram reutilizadas 13 vezes em páginas da Wikipédia em português e 19 vezes em 7 páginas de Wikipédias de outros idiomas. Das categorias selecionadas para a pesquisa foi a que apresentou maior ocorrência de imagens, ilustrando itens no Wikidata com 4 ocorrências. Das 471 imagens disponibilizadas na categoria, há 9 imagens sendo reutilizadas em 36 ocorrências, o que constitui 2,55% de reuso do total de imagens disponíveis na categoria.

A coleção referente ao Museu Castro Maya é a categoria que apresenta maior quantidade de imagens distintas utilizadas. A categoria disponibiliza o total de 46 imagens, das quais 17 imagens indicaram ocorrência em 12 páginas da Wikipédia em português, em 25 páginas de Wikipédias de outros idiomas, em cinco páginas do Wikibooks e em três itens do Wikidata, somando assim 45 ocorrências de utilização de imagens, o que corresponde ao um reuso de 36,96% de todas as imagens disponíveis na categoria.

Apesar de apontar somente sete imagens na categoria referente ao Museu da Inconfidência, essa coleção é a que apresenta a maior porcentagem de ocorrência de usos das imagens disponíveis, com o uso de 6 imagens distintas em 8 páginas da Wikipédia em português e em 10 artigos de outros idiomas, resultando na ocorrência do total de oito reusos de imagens.

Tabela 1 – Reuso de imagens nos projetos da Fundação Wikimedia

Coleção	Arquivos na categoria	Wikipédia português	Wikipédia em outros idiomas	Wikibooks	Wikiquote	Wikidata
Museu da Inconfidência	7	8	10	0	0	0
Museu da República	9	33	75	1	1	0
Museu Nacional de Belas Artes	17	2	4	0	0	1
Museus Castro Maya	46	12	25	5	0	3
Museu Casa de Benjamin Constant	31	4	0	0	0	0
Museu Regional de São João del-Rei ⁴	6	0	0	0	0	0
Museu Victor Meirelles	4	5	12	0	0	3
Museu Histórico Nacional	471	13	19	0	0	4
Museu Imperial	38	7	52	0	0	0

Fonte: Dados da pesquisa, 2020⁵.

Tabela 2 – Quantidade de imagens reutilizadas nos projetos da Fundação Wikimedia

Coleção	Arquivos na categoria	Total de imagens distintas	Total de uso de imagens	Imagens da categoria (%)
Museu da Inconfidência	7	6	18	85,71
Museu da República	9	4	110	44,44
Museu Nacional de Belas Artes	17	7	7	17,65
Museus Castro Maya	46	17	45	36,96
Museu Casa de Benjamin Constant	31	4	4	12,9
Museu Regional de São João del-Rei	6	0	0	0
Museu Victor Meirelles	4	3	20	75
Museu Histórico Nacional	471	12	36	2,55
Museu Imperial	38	8	59	21,05

Fonte: Dados da pesquisa, 2020.

⁴ Apesar de identificada a categoria com 6 mídias referentes à coleção ao Museu de São João del-Rei, nenhuma das ferramentas utilizadas retornou dados sobre ela.

⁵ Os dados da pesquisa podem ser acessados em: https://docs.google.com/spreadsheets/d/1BNYn2WKGgqOT0UvC_Nla9OgciK3I8V_6NRfg5ckE4-8/edit?usp=sharing

A REUTILIZAÇÃO DAS IMAGENS DAS CATEGORIAS REFERENTE ÀS COLEÇÕES DE MUSEUS NA WIKIPÉDIA

A Wikipédia é um projeto de enciclopédia online de edição coletiva e colaborativa da Fundação Wikimedia. Devido à sua proposta multilíngue, atualmente existem versões da Wikipédia em 303 idiomas. Cada idioma se mostra como um projeto Wikipédia independente do outro, ou seja, os conteúdos não são automaticamente traduzidos e reproduzidos, sendo originalmente criados e geridos pelos usuários de cada comunidade linguística e apresentam seu subdomínio. A Wikipédia, em seus variados idiomas, tem seus conteúdos baseados na licença aberta Creative Commons Attribution-ShareAlike.

Atualmente⁶ a Wikipédia lusófona alcança a marca de 1.029.760 artigos produzidos, 6.110 usuários ativos (WIKIPEDIA, 2019).

Para entender melhor como se dá a reutilização dessas imagens tanto em páginas da Wikipédia lusófona, quanto em outros idiomas, buscamos coletar dados sobre a imagem da categoria referente ao Museu da República que foi reutilizada com maior frequência⁷. Por meio da ferramenta GLAMorous verificamos que a imagem com maior utilização é uma fotografia do ex-presidente brasileiro Prudente de Moraes. Esse arquivo intitulado “Prudentedemoraais.jpg” aparece com 95 ocorrências de reutilização.

Figura 1 – Retrato de Prudente de Moraes na Wikimedia Commons



Fonte: Wikimedia Commons, 2020⁸.

⁶ Dados coletados em abril de 2020.

⁷ Em abril de 2020.

⁸ Disponível em: <<https://commons.wikimedia.org/wiki/File:Prudentedemoraais.jpg>>. Acesso em abr. 2020.

Ao observar a imagem no contexto da plataforma Wikimedia Commons, é possível verificar que a página da imagem oferece meios de download e compartilhamento da imagem, e também exibe uma série de seções com informações da mídia. A seção intitulada “Informações do arquivo” fornece uma série de informações sobre a imagem como “Descrição”, “Data”, “Permissão”⁹ e “Outras versões”.

Além disso, podem ser observadas outras seções como “Histórico do arquivo”, que mostra o histórico de substituição de arquivos no Wikimedia Commons, a seção “Uso de arquivos no Commons”, que dá informações sobre o uso de versões no Commons, a seção “Uso de arquivos em outras wikis”, que gera uma lista de projetos wiki nos quais a imagem está sendo utilizada, e por fim a seção “Metadado”, que reúne metadados e dados específicos da mídia reprodutora da imagem, que nesse caso são específicos à fotografia que resultou na imagem digitalizada.

Ao consultar a listagem na seção “Uso de arquivos de outras wikis”, é possível observar em quais outros idiomas de Wikipédias o retrato de Prudente de Moraes está sendo utilizado e em quais páginas específicas. Essas informações também podem ser obtidas por meio de consultas à ferramenta GLAMorous.

Assim foi possível verificar a presença do retrato de Prudente de Moraes em artigos de Wikipédias em 44 idiomas distintos e também em uma página da Wikiquote em português. Também foi possível verificar que a mesma imagem está presente em 30 artigos da Wikipédia em português. No quadro 2 é possível verificar a listagem desses artigos.

Quadro 2 – Lista de artigos que reutilizam a imagem o retrato de Prudente de Moraes

pt.wikipédia	Prudente de Moraes Rodrigues Alves Presidente Prudente Lista de eleições presidenciais no Brasil Presidente do Brasil Bernardino José de Campos Júnior Manuel José Alves Barbosa Eleição presidencial no Brasil em 1891 Amaro Cavalcanti Manuel Vitorino Joaquim Murtinho Antônio Olinto dos Santos Pires Antônio Gonçalves Ferreira Dionísio Evangelista de Castro Cerqueira Alberto Torres Carlos Machado de Bittencourt Carlos Augusto de Carvalho Lista de presidentes do Senado Federal do Brasil João Tomás de Cantuária Jerônimo Rodrigues de Moraes Jardim Sebastião Eurico Gonçalves de Lacerda Francisco de Paula Argolo Bernardo Vasques Elisiário José Barbosa Eleição presidencial no Brasil em 1894 Lista de presidentes do Brasil Governo Prudente de Moraes Lista de senadores do Brasil da 23.ª legislatura Lista de senadores do Brasil da 22.ª legislatura
--------------	---

Fonte: Dados da pesquisa, 2020.

Percebe-se que a imagem é reutilizada sobretudo para ilustrar verbetes de personagens históricos que tiveram relação com o ex-presidente ou eventos importantes relacionados a ele. Além disso, aparecem na lista alguns eventos históricos, listas de senadores em legislaturas específicas, eleições presidenciais e o próprio verbete relativo ao seu governo. A imagem do acervo do museu é colocada em contexto a partir da textualidade dos verbetes e das relações simbólicas que estabelece com outras imagens e elementos textuais. Estudar essas relações disponibiliza uma camada informacional de compreensão do reuso da informação a ser desenvolvida em pesquisas posteriores, mas é algo a ser destacar como potencial para a área.

⁹ Permissões de reuso do artigo, especificando as licenças atribuídas.

A REUTILIZAÇÃO DE IMAGENS NO REPOSITÓRIO DE DADOS ESTRUTURADOS WIKIDATA

Outro aspecto interessante a ser verificado é a reutilização das imagens de coleções do universo de museus e instituições culturais na ilustração de itens do Wikidata. O Wikidata é o projeto de repositório de dados estruturados da Fundação Wikimedia e tem “como objetivo a criação de uma base de conhecimento livre sobre o mundo que pode ser lida e editada por humanos e máquinas” (WIKIDATA, 2019). Segundo dados coletados em abril de 2020 na página do projeto, o Wikidata registrava 83.779.497 milhões de itens em seu banco, 26.160 usuários ativos no mês da coleta e 308 *bots* que auxiliam nas edições. Os dados publicados no Wikidata também são disponibilizados de forma aberta sob a licença Creative Commons Public Domain Dedication 1.0.

As regras de criação e gerenciamento do conteúdo são de responsabilidade da comunidade de editores, que devem seguir os mesmos princípios e políticas de colaboração que norteiam os demais projetos Wikimedia.

Para visualizar e entender a reutilização de imagens do Wikimedia Commons no Wikidata, selecionamos para a análise a coleção que apresentou, na tabela 1, o número maior de ocorrências de uso de imagens no Wikidata, que foi o caso do Museu Histórico Nacional. Usando a ferramenta GLAMorous e explorando os resultados detalhados retornados por meio de busca com o nome da categoria, foi possível verificar quais imagens foram utilizadas nas 4 ocorrências detectadas e também o código identificador do item no Wikidata, como pode ser visto no quadro 3.

Quadro 3 – Imagens reutilizadas na Wikidata

Nome do arquivo	ID Wikidata
Conselheiro Francisco de Carvalho Soares Brandão, da coleção Museu Histórico Nacional.jpg	Q62091980
Beatriz Reinall, da coleção Museu Histórico Nacional.jpg	Q62091982
Caminho na floresta, da coleção Museu Histórico Nacional.jpg	Q62091983
Viscondessa do Bom Conselho, da coleção Museu Histórico Nacional.jpg	Q62091547

Fonte: Dados da pesquisa, 2020.

Figura 2 – Retrato de Prudente de Moraes na Wikimedia Commons.

Fonte: Wikimedia Commons, 2020¹⁰.

¹⁰ Disponível em: < <https://www.wikidata.org/wiki/Q62091980>>. Acesso em 15 de abr. 2020.

Para visualizar como se dá essa reutilização no contexto da plataforma Wikidata, realizamos buscas na plataforma por meio do identificador do arquivo da imagem intitulada “Conselheiro Francisco de Carvalho Soares Brandão, da coleção Museu Histórico Nacional.jpg”. Desse modo foi possível observar que a mídia foi inserida na propriedade “imagem” do item no Wikidata, que apresenta dados estruturados sobre a obra cuja original analógica está sob guarda no acervo do Museu Histórico Nacional.

Ao verificar a reutilização do arquivo intitulado “Beatriz Reinall, da coleção Museu Histórico Nacional.jpg”, foi possível identificar que essa imagem também ilustra um item do Wikidata que apresenta dados estruturados sobre a obra de arte retratada. A mesma situação se ocorre no caso das outras duas imagens identificadas que também ilustram itens sobre as obras originais. Com essas informações, é possível inferir que foi realizado um trabalho coordenado de descrição de obras do Museu Histórico Nacional na Wikidata. Em uma investigação realizada em 2019, os autores Carmo e Martins (2019, p. 12) identificaram que o Museu Histórico Nacional representava o segundo maior acervo oriundo de instituições brasileiras com itens de pinturas descritas no Wikidata, reunindo 503 itens. Esses dados indicam que a instituição estabeleceu em algum momento, ou ainda estabelece, práticas que adotam as plataformas da Fundação Wikimedia como meios estratégicos de disseminação de seus acervos.

A VISUALIZAÇÃO DE ARQUIVOS DE MÍDIA DO WIKIMEDIA COMMONS

Utilizando a ferramenta GLAMorous 2, foi possível obter dados de visualização dos arquivos de mídias disponibilizados na Wikimedia Commons que estão nas categorias referentes às coleções de museus geridos pelo Ibram. Essas visualizações são contabilizadas com base nos dados de acesso a páginas de projetos da Fundação Wikimedia, como o Wikidata, Wikibooks e Wikiquote, e também das Wikipédias dos mais variados idiomas que exibem imagens do Wikimedia Commons reutilizadas em suas páginas. O recorte utilizado alcança dados de visualização relativos aos 12 meses do ano de 2019, de janeiro a dezembro. Na tabela 3 podemos verificar o total de visualizações de arquivos de mídia no ano de 2019, assim como a média mensal. Também foi possível observar o total de visualizações somente na Wikipédia em português, assim como o total de visualizações em outras Wikipédias e projetos.

Com base nos dados coletados, é possível observar alguns números que nos chamam especial atenção. Tomamos como exemplo os dados de visualização da coleção referente ao Museu da República, que com apenas 9 arquivos na categoria foi a que gerou o número de visualizações de arquivos de mídias mais expressivo no ano de 2019, alcançando a marca de 4.147.494 de visualizações.

Tabela 3 – Reuso de imagens nos projetos da Fundação Wikimedia

Coleção	Total de visitas - 2019	Média mensal	Total de visualizações pt wikipédia	Total de visualizações outras wikis
Museu da Inconfidência	675.728	56.310,67	635.523	40.205
Museu da República	4.147.494	345.624,50	3.225.730	921.764
Museu Nacional de Belas Artes	21.873	1.822,75	20.904	969
Museus Castro Maya	882.608	73.550,67	795.007	87.601
Museu Casa de Benjamin Constant	6.216	518,00	6.216	0
Museu Regional de São João del-Rei	0	-	0	0
Museu Victor Meirelles	69.089	5.757,42	44.791	24.298
Museu Histórico Nacional	578.668	48.222,33	45.643	45.643
Museu Imperial	650.276	54.189,67	174.333	475.943

Fonte: Dados da pesquisa, 2020.

Ao recorrer à opção “Detalhes de uso do arquivo” na ferramenta GLAMorous2, é possível observar que a imagem identificada como “Prudentedemorais.jpg”¹¹, já anteriormente identificada como a mídia que mais apresenta reúsos dentro da coleção, revela-se sozinha como a responsável por 3.096.852 do total de visualizações das mídias que estão na categoria da coleção do museu. Seguida da imagem intitulada “GetulioVargasPijamaRevolver.jpg”, que registrou no ano de 2019 o total de 992.687 visualizações. Também foi possível constatar que 3.225.730 visualizações das mídias da coleção do Museu da República foram somente em Wikipédias em português, e 675.728 em Wikipédias de outros idiomas e em outros projetos da Wikimedia. A título de ilustração, para que se possa problematizar o fenômeno social em questão, no primeiro semestre do ano de 2019 o Museu da República teve em torno de 105.000 visitantes presenciais, segundo dados fornecidos pelo próprio Instituto Brasileiro de Museus (BRASIL, 2019). Se tomarmos por partido a comparação com os números digitais, temos que o número de pessoas que visualizaram objetos do acervo presencialmente representa em torno de 2,5% dos que visualizaram os objetos digitais. Sem dúvida, e é fundamental destacar esse ponto, não se quer com isso fazer uma equivalência das experiências, sabendo que as visitas presenciais são mediadas de diversos outros elementos culturais que precisam ser reconhecidos e que possuem efeitos importantes na apropriação. Mas compreender esse espaço de socialização e sua dimensão simbólica como estratégia complementar para o campo dos museus é algo de enorme importância nos tempos atuais. Há um espaço de trabalho que pode ser aprimorado e que contém largo potencial de socialização e produção de valor social pelo museu.

Outro dado interessante foi a quantidade de visualizações de mídias referente a categoria do Museu Imperial, que apresentou maior número de visualizações em Wikipédias de outros idiomas, com 475.943 visualizações, do que na Wikipédia em português, com 174.333 visualizações. Ao verificar esses dados na aba da ferramenta chamada “Uso global do arquivo”, foi possível constatar que só a Wikipédia em inglês indica o total de 319.974 visualizações e faz 14 usos de 7 diferentes mídias da coleção do Museu Imperial em suas páginas¹², no ano 2019.

Já as mídias da categoria referente ao Museu Benjamin Constant chamam atenção por não apontarem visualizações em outros projetos da Wikimedia ou em Wikipédias de outros idiomas, somente na Wikipédia em português com 6.216 visualizações no ano de 2019, representando a coleção que menos obteve visualizações dentro do universo investigado.

CONCLUSÃO

O presente artigo trouxe resultados de uma investigação que buscou explorar possibilidades de coleta de dados que permitissem a realização de análises webométricas que permitem mensurar aspectos relacionados ao acesso e à reutilização de mídias de acervos brasileiros disponíveis na internet, por meio do repositório de mídias Wikimedia Commons. Assim, verificamos a presença de coleções de nove dos 30 museus sob responsabilidade do Ibram entre as categorias do Wikimedia Commons. Identificadas essas categorias e com o auxílio das ferramentas apresentadas, obtivemos e descrevemos números relativos à reutilização dessas mídias tanto em páginas da Wikipédia em português, das Wikipédias em outros idiomas, quanto em outros projetos da Fundação Wikimedia.

¹¹ A ferramenta GLAMorous 2 apresenta também a estatística de uso da imagem “Prudentemorais.jpg”, que indica 106 usos. Ao comparar com os dados obtidos pela GLAMorous, foi possível observar uma discrepância, já que ela nos apresenta um número diferente de usos para essa imagem, neste caso 95 usos. Ao verificar os dados e mais especificamente a lista que especifica as páginas, percebemos que a ferramenta GLAMorous 2 conta todas as páginas dos projetos, incluindo páginas de usuários que não são artigos, enquanto a GLAMorous mostra somente os usos em artigos.

¹² Esses dados podem ser verificados no link de consulta: <<https://tools.wmflabs.org/glamtools/glamorous/?mode=category&text=Collections%20of%20the%20Museu%20Imperial&pageviews=1&month=2019-01%7C2019-12>> Acesso em 15 de abril, 2020.

Assim foi possível observar que o número de mídias disponíveis não afeta diretamente o número de reutilizações ou de visualizações, já que o Museu da República se destacou tanto em número de reusos quanto em números de visualizações, com apenas 9 arquivos na categoria. Apesar de o Museu Histórico Nacional conter o maior número de mídias, 471 arquivos, não obteve números relevantes nos dados de reuso ou de visualização deles. Também foi possível observar que a mesma imagem de determinada coleção pode ser utilizada diversas vezes, gerando uma porcentagem baixa do total de uso dos arquivos. Outra situação que se repetiu foi a baixa reutilização de mídias para ilustrar outros projetos da Fundação Wikimedia, como o Wikidata. A presença de mídias e a quantidade de visualizações nas Wikipédias de diferentes idiomas foram dados que nos chamaram atenção, com ênfase no Museu da República, e mais especificamente no caso do retrato do ex-presidente brasileiro Prudente de Moraes, que aparece ilustrando 30 artigos da Wikipédia em português e 45 artigos de Wikipédias de outros idiomas. Outro dado curioso a ser mencionado foi o fato de o Museu Imperial apresentar o maior número de visualizações na Wikipédia em inglês do que na Wikipédia em português. Por fim, outro dado relevante foi o número de visualizações obtidos pelos arquivos na categoria referente a coleções do Museu da República, que concentrou em apenas uma imagem mais de 3 milhões de visualizações, somente no ano de 2019. Durante a investigação foi possível perceber evidências que apontam para a expressividade do público digital em relação ao público presencial, deixando elementos que podem ser explorados em pesquisas posteriores, procurando qualificar e entender que público é esse, assim como os impactos que esse tipo de visita pode gerar para os museus e como isso pode refletir nas práticas de seus profissionais.

Por meio dos dados obtidos e das análises realizadas foi possível observar que a plataforma Wikimedia Commons, assim como outras plataformas da Fundação Wikimedia, como a Wikipédia e Wikidata, são espaços sociotécnicos que além de se configurar como meios potenciais de socialização de acervos, também se configuram como fonte de informação valiosa sobre a manipulação e uso das informações disponibilizadas por meio deles. Acreditamos que a obtenção desse tipo de dado pode ajudar instituições culturais guardiãs de acervos, como os museus, a conhecer e a compreender o que acontece com seus acervos digitais uma vez que são disponibilizados nesses espaços, fornecendo importantes informações para a gestão das coleções de seu acervo digital, principalmente em relação ao acesso e reapropriação desses acervos por parte dos usuários.

REFERÊNCIAS

- BJÖRNEBORN, L.; INGWERSEN, P. Toward a basic framework for webometrics. *Journal of the American Society for Information Science and Technology*, v. 55, n. 14, p. 1216–1227, dez. 2004. DOI 10.1002/asi.20077. Disponível em: <http://doi.wiley.com/10.1002/asi.20077>. Acesso em: 12 mar. 2021.
- BRASIL. MINISTÉRIO DA CIDADANIA. *Plano Nacional de Cultura. Museus brasileiros apresentam aumento no número de visitantes*. 19 ago. 2019. Disponível em: <http://pnc.cultura.gov.br/2019/08/19/museus-brasileiros-apresentam-aumento-no-numero-de-visitantes/>. Acesso em: 30 abr. 2020.
- CARMO, D. do; MARTINS, D. L. A presença dos museus brasileiros na ecologia informacional da Fundação Wikimedia: estudo de caso do projeto Sum of All Paintings. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 20., 2019. *Anais [...]*. Florianópolis: UFSC, 2019. v. 20, p. 1–20. Disponível em: <https://conferencias.ufsc.br/index.php/enancib/2019/paper/view/909/679>. Acesso em: 12 mar. 2021.
- INSTITUTO BRASILEIRO DE MUSEUS. *Museus IBRAM*. 2020. Disponível em: <https://www.museus.gov.br/os-museus/museus-IBRAM/>. Acesso em: 5 abr. 2020.
- MANSKE, M. *GLAMorous*. 2020a. Disponível em: <https://tools.wmflabs.org/glamtools/glamorous.php>. Acesso em: 18 abr. 2020.
- MANSKE, M. *GLAMorous 2*. 2020b. Disponível em: <https://tools.wmflabs.org/glamtools/glamorous/>. Acesso em: 18 abr. 2020.

STINSON, A. D.; WYATT, L.; FAUCONNIER, S. Stepping Beyond Libraries : the Changing Orientation in Global GLAM-Wiki. *JLIS*, n. 9, 2018. DOI 10.4403/jlis.it-12480. Disponível em: <https://doi.org/10.4403/jlis.it-12480>. Acesso em: 12 mar. 2021.

THELWALL, M. Introduction to Webometrics: Quantitative Web Research for the Social Sciences. *Synthesis Lectures on Information Concepts, Retrieval, and Services*, v. 1, n. 1, p. 1–116, jan. 2009. DOI 10.2200/S00176ED1V01Y200903ICR004. Disponível em: <http://www.morganclaypool.com/doi/abs/10.2200/S00176ED1V01Y200903ICR004>. Acesso em: 12 mar. 2021.

VANTI, N. A. P. Da bibliometria à webometria: uma exploração conceitual dos mecanismos utilizados para medir o registro da informação e a difusão do conhecimento. *Ciência da Informação*, v. 31, n. 2, p. 369–379, ago. 2002. DOI 10.1590/S0100-19652002000200016. Disponível em: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-19652002000200016&lng=pt&tlng=pt. Acesso em: 12 mar. 2021.

VILLAESPESA, E.; NAVARRETE, T. Museum Collections on Wikipedia: Opening Up to Open Data Initiatives. In: MUSEWEB, 14 jan. 2019. *Proceedings* [...]. Boston: [s. n.], 14 jan. 2019. Disponível em: <https://mw19.mwconf.org/paper/museum-collections-on-wikipedia-opening-up-to-open-data-initiatives/>. Acesso em: 12 mar. 2021.

WIKIMEDIA COMMONS. *Commons: Welcome*. 2020a. Disponível em: <https://commons.wikimedia.org/wiki/Commons:Welcome>. Acesso em: 15 abr. 2020.

WIKIMEDIA COMMONS. *Special: Statistics*. 2020b. Disponível em: <https://commons.wikimedia.org/wiki/Special:Statistics>. Acesso em: 18 abr. 2020.

WIKIMEDIA OUTREACH. *GLAM/Resources/Tools*. 2020. Disponível em: <https://outreach.wikimedia.org/wiki/GLAM/Resources/Tools>. Acesso em: 20 mar. 2020.

WIKIPEDIA. *Página Principal*. 2019. Disponível em: https://pt.wikipedia.org/wiki/Wikip%C3%A9dia:P%C3%A1gina_principal. Acesso em: 18 out. 2020.

ZEINSTRAS, M. *Report on Requirements for Usage and Reuse Statistics for GLAM Content*. [S. l.]: Wikimedia, 31 maio 2013. Disponível em: <https://www.kl.nl/en/publications/report-on-requirements-for-usage-and-reuse-statistics-for-glam-co/>. Acesso em: 20 set. 2019.