

**Publicação, difusão e colaboração em rede na documentação de acervos digitais de instituições de memória: a interoperabilidade entre as redes de informação Wikidata, Wikimedia Commons, Wikipédia e o software livre Tainacan**

Dalton Lopes Martins  
Universidade de Brasília  
Faculdade de Ciência da Informação

Resumo: O projeto propõe o desenvolvimento de um serviço que facilite a publicação e o monitoramento das edições e colaborações realizadas por usuários de acervos digitais de instituições de memória nos ambientes Wikidata, Wikimedia Commons e Wikipédia. Para tanto, propõe a modelagem técnica de um serviço de retroalimentação, denominado *roundtripping*, entre as redes de informação do ecossistema wiki e o software livre Tainacan. O Tainacan vem se configurando como um importante software livre para a gestão e a difusão de acervos digitais em rede de instituições de memória no Brasil. Existem consideráveis desafios técnicos e operacionais para a implementação de um serviço que não só permita a publicação de acervos a partir de um repositório digital de uma instituição de memória, mas também o monitoramento do reuso, edições e colaborações de usuários em outras redes de informação. Ao permitir a conexão entre as redes de informação do Wikidata, do Wikimedia Commons e do Wikipédia para a publicação e o estímulo ao reuso dos acervos digitais das instituições já publicadas no Tainacan, o projeto visa ampliar a circulação das coleções patrimoniais, tornar relevante o conhecimento gerado sobre elas e, assim, valorizar a cultura material brasileira na sociedade em rede.

**Publication, dissemination and network collaboration in the documentation of digital collections of memory institutions: interoperability between the information environments Wikidata, Wikimedia Commons, Wikipedia and the free software Tainacan**

Dalton Lopes Martins  
University of Brasilia  
Faculty of Information Science

Abstract: The project proposes the development of a service that facilitates the publication and monitoring of editions and collaborations carried out by users of digital collections from memory institutions in the Wikidata, Wikimedia Commons and Wikipedia environments. Therefore, it proposes the technical modeling of a feedback service, called roundtripping, between the information networks of the wiki ecosystem and the free software Tainacan. Tainacan has become an important free software for the management and dissemination of digital collections in a network of memory institutions in Brazil. There are important technical and operational challenges for the implementation of a service that not only allows the publication of collections from a digital repository of a memory institution, but also monitors the reuse, editions and collaborations of users in other information networks. By allowing the connection between the Wikidata, Wikimedia Commons and Wikipedia information networks for the publication and encouragement of the reuse of the digital collections of institutions already published in Tainacan, the project aims to expand the circulation of heritage collections, make the knowledge generated about them relevant and, thus, value Brazilian material culture on the network society.

## 1. Enunciado do problema

As redes de informação Wikipédia, Wikidata e Wikimedia Commons constituem-se, atualmente, como um conjunto de importantes recursos informacionais disponíveis na Internet gratuitamente para o público, sendo amplamente utilizados pela sociedade. A importância desses recursos pode ser melhor compreendida quando os números relativos ao seu funcionamento são explicitados. A Wikipédia, por exemplo, funciona como uma enciclopédia aberta à colaboração de qualquer usuário, possuindo políticas e regras de governança que instauram dinâmicas de moderação coletiva, ampliando a qualidade e o potencial informativo dos conteúdos produzidos. Atualmente disponível em 310 línguas ativas, com mais de 56 milhões de artigos produzidos, mais de 96 milhões de usuários e mais de 3.800 administradores em termos mundiais (WIKIMEDIA FOUNDATION, 2021b), a Wikipédia se constitui como uma fonte de informação diversa e abrangente, aparecendo como uma das primeiras páginas em buscadores, como o Google, para diversos temas de interesse. O acesso dos usuários à Wikipédia se dá, em 65% dos casos, a partir de uma busca no Google, no qual o mecanismo de busca apresenta um link de referências nas primeiras páginas ou em suas caixas de informação laterais, tornando-se assim um dos 5 sites mais visitados do mundo (NAVARRETE e VILLAESPESA, 2020). A título de exemplo, no último ano (dados de junho de 2020 a maio de 2021) apenas a Wikipédia em língua portuguesa (<https://pt.wikipedia.org/>) teve mais de 391 milhões de visualizações e 228 mil edições realizadas por aproximadamente 3 mil editores ativos na plataforma (WIKIMEDIA FOUNDATION, 2021a).

O Wikimedia Commons, por sua vez, funciona como um repositório de arquivos de mídia licenciados de maneira a permitir o reuso livre, pelos usuários em outros sites e projetos na web, contando atualmente com mais de 74 milhões de arquivos publicados (WIKIMEDIA COMMONS, 2021). Por fim, o Wikidata funciona como uma base de conhecimento livre, permitindo a estruturação de dados semânticos por meio de declarações em formato de triplas sujeito-predicado-objeto, contando atualmente com mais de 93 milhões de itens publicados e se tornando um nó central para a viabilização da web semântica na Internet (WIKIDATA, 2021).

Os 3 serviços juntos se integram de maneira a proverem um ecossistema de circulação e produção colaborativa da informação. Entre suas possibilidades de integração, destacam-se as três descritas a seguir. As imagens publicadas no Commons podem ser utilizadas para ilustrar verbetes publicados na Wikipédia, auxiliando na oferta de recursos informacionais de caráter pedagógico para a melhor compreensão dos usuários. Os dados estruturados no Wikidata podem ser facilmente recuperados e serem apresentados como caixas de informação nos

verbetes da Wikipédia, facilitando o consumo de dados quantitativos, temporais e categóricos. Os arquivos publicados no Commons podem ser descritos em seus aspectos de mídia, de conteúdo e de proveniência a partir de metadados organizados em modelos semânticos que se integram no Wikidata, facilitando o processo de catalogação e ampliando o potencial de busca, recuperação e reuso por outros sistemas de informação.

Algumas pesquisas recentes vêm demonstrando que instituições de memória como arquivos, bibliotecas e museus têm se tornado importantes provedores de dados para o ecossistema wiki, obtendo resultados de impacto considerável ao difundir seus acervos nessas redes de informação. Villaespesa e Navarrete (2019), pesquisando na Wikipédia em língua inglesa, identificaram que 8.104 pinturas oriundas de 785 museus de 59 países do mundo todo foram utilizadas em 10.008 artigos na Wikipédia. Esses artigos juntos tiveram em média mais de 94 milhões de visualizações mensais durante o ano de 2017. Em uma segunda pesquisa, Navarrete e Villaespesa (2020) identificaram no Wikidata 224.374 itens catalogados como pintura, sendo que destes 89.637 (40%) possuíam metadados e 27.501 (12%) possuíam imagem. Ao estudar o uso dessas imagens fora do contexto de páginas e artigos específicos do mundo das artes, as pesquisadoras identificaram que os 3 temas que mais receberam pinturas oriundas de museus para ilustrar verbetes na Wikipédia foram "história" (3.034 pinturas), "religião" (924 pinturas) e "geografia" (600 pinturas). O estudo busca mostrar como as coleções digitais de museus publicadas tornam-se provedoras de dados para diversas temáticas e apoiam a construção dos verbetes. Ferriter (2019) relata o estudo de caso focado no Wikidata iniciado pela Biblioteca do Congresso Americano, no qual foram identificadas mais de 650.000 referências para os identificadores oriundos dos sistemas de controle de autoridade mantidos pela biblioteca, a saber, o *Name Authority File* (NAF) e o *Library of Congress Subject Headings* (LCSH). Durante a execução do caso, a biblioteca realizou um experimento no qual o usuário podia navegar em mais de 66 mil imagens, que representavam 13.300 entidades descritas por metadados bibliográficos oriundos do Wikidata. A Association of Research Libraries (2019) publicou um relatório de recomendações e oportunidades para a adoção do Wikidata por bibliotecas. Um dos principais destaques que o relatório menciona em sua conclusão é a possibilidade de utilização do Wikidata para a integração de vários sistemas, maximizando seu potencial de interoperabilidade a partir de seu papel em fornecer significado à organização da informação de diferentes estratégias de controle de autoridades.

No Brasil, Martins e Carmo (2019) estudaram como as páginas dos museus ligados ao Instituto Brasileiro de Museus (IBRAM) foram construídas na Wikipédia. Foram identificadas na pesquisa páginas de 20 museus editadas por 555 colaboradores que realizaram 1.108

edições. A pesquisa aponta uma dinâmica de construção coletiva de informações em rede sobre museus, evidenciando um agenciamento expressivo de colaboradores na construção de informações de interesse público. Carmo e Martins (2019b) também estudaram a presença de museus brasileiros no projeto liderado pela comunidade que colabora nos projetos da Fundação Wikimedia denominado "*Sum of all paintings*", projeto que tem por objetivo reunir dados de pinturas em âmbito mundial no Wikidata. A pesquisa identificou dados oriundos de 30 países e 56 colaboradores. A pesquisa focou em identificar a presença de instituições brasileiras nesse ambiente, encontrando 19 coleções de 14 museus, representando um total de 3.583 pinturas. Carmo e Martins (2019a) também estudaram a presença de acervo dos museus ligados ao Instituto Brasileiro de Museus no Wikimedia Commons. Nesse estudo, identificaram que 9 desses museus estavam presentes com 629 imagens de pinturas associadas. Chama a atenção na pesquisa o total de visualizações das imagens publicadas pelo Museu da República, que tiveram mais de 4 milhões de visitas em 2019. As imagens do acervo desse museu foram utilizadas para ilustrar verbetes históricos de grande interesse público como, por exemplo, a página do presidente Prudente de Moraes na Wikipédia. Oliveira e Martins (2020) realizam um estudo exploratório para reconciliação de metadados oriundos do acervo digital do Museu Histórico Nacional com o Wikidata. Os pesquisadores demonstram que dos metadados escolhidos apenas o "autor" obteve resultados significados de reconciliação, evidenciando a falta de vocabulários do patrimônio cultural brasileiro na plataforma Wikidata. As pesquisas mencionadas e os números relatados mostram o potencial para a circulação e a produção colaborativa da informação em rede nos ambientes wiki. As pesquisas ressaltam ainda a baixa participação de instituições culturais brasileiras<sup>1</sup>, o que permite inferir um potencial ainda a ser explorado.

Entende-se, a partir do exposto, que a presença dos acervos oriundos das instituições culturais no ecossistema wiki gera um ciclo virtuoso de reutilização de dados em contextos variados, ampliando seu potencial de apropriação cultural. Além disso, observa-se o potencial da retroalimentação dos dados oriundos da catalogação original dos objetos a partir das sugestões, correções e revisões realizadas pelos usuários dessas redes integradas de conhecimento. Como ressalta Monteiro (2019, p.72):

Os projetos Wikimedia colaboram na criação de novos links de hipertexto para tal conteúdo, permitindo que eles sejam reeditados, remixados e colocados em novas lógicas interpretativas. Em outras palavras, é possível dizer que as

---

<sup>1</sup> [https://pt.wikipedia.org/wiki/Wikip%C3%A9dia:GLAM/Projetos\\_em\\_portugu%C3%AAs](https://pt.wikipedia.org/wiki/Wikip%C3%A9dia:GLAM/Projetos_em_portugu%C3%AAs). Acesso em 16 jun. 2021.

parcerias podem viabilizar o compartilhamento de coleções culturais abertamente por meio de plataformas populares e conhecidas, como a Wikipédia. E os museus só podem desfrutar de todos os benefícios dessa abertura para o mundo.

Uma vez compreendido o potencial social e cultural da presença dos acervos das instituições culturais no ecossistema wiki, cabe perguntar como esses acervos deveriam ser publicados nesse ambiente e, especialmente, como as instituições poderiam acompanhar os seus objetos publicados, tendo a capacidade de receber e avaliar a pertinência de sugestões de melhorias, correções ou revisões na documentação do objeto por usuários e robôs do Wikidata e Wikimedia Commons, bem como monitorar o número de visualizações e os verbetes nos quais suas imagens são utilizadas na Wikipédia. Segundo Villaespesa e Navarrete (2019), a chave para tratar esse problema de forma eficiente é integrar a solução técnica no fluxo de trabalho organizacional da documentação das instituições de memória. Para isso, as autoras propõem o fluxo apresentado na figura 01. Observa-se, destacado em vermelho, a base de dados do museu a partir de onde, à direita, parte um segmento cuja ramificação leva à seta de publicação de metadados no Wikidata e outra de publicação da mídia relacionada ao item no Wikimedia Commons. Esses dados, uma vez publicados, podem ser utilizados por usuários na Wikipédia para ilustrarem verbetes, como visto na parte de cima da figura destacada em azul. É interessante notar que retorna para o museu, com origem no ecossistema wiki, informações relacionadas a traduções, melhorias nos dados e ferramentas exploratórias.

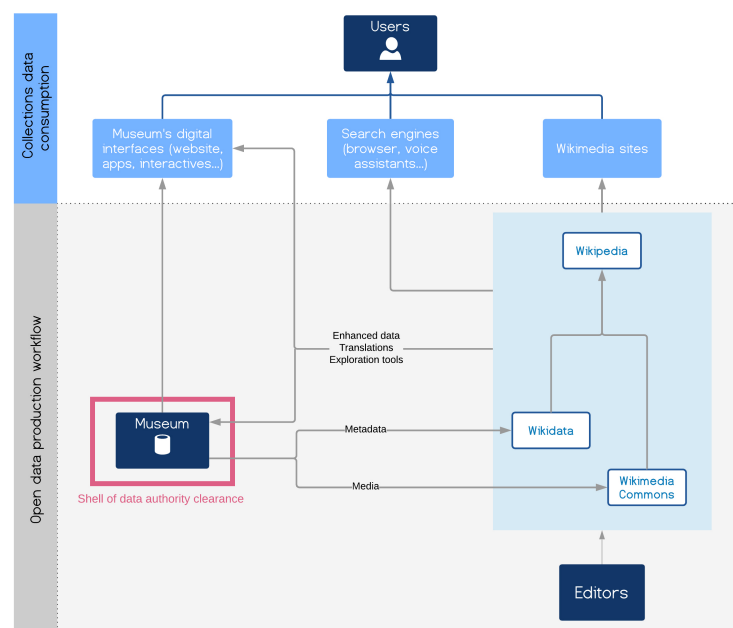


Figura 01. Fluxo de produção e retroalimentação de dados entre sistema de um museu e ecossistema wiki. Fonte: (VILLAESPESA e NAVARRETE, 2019).

O elemento com maior destaque na figura 01, e que parece ser a chave do processo de automação do fluxo operacional de publicação e retroalimentação de dados do ecossistema wiki, é o sistema de base de dados do museu. Presume-se que esse sistema tenha a possibilidade técnica, a partir de padrões de anotação associados, de interoperar com o Wikidata e o Wikimedia Commons para que tal fluxo se torne de fato operacional e automatizado. Tal abordagem assegura que o significado intencionado da semântica ligada aos dados das bases dos museus possa ser compartilhado entre diferentes aplicações no escopo do ecossistema wiki. Além do compartilhamento semântico, o modelo conceitual ligado às bases de dados dos museus deve prever meios de transmissão em alguma sintaxe acordada pela comunidade do Wikidata, que, neste caso, seria por meio de formatos compatíveis com a infraestrutura da web semântica como RDF/OWL (HENDLER; GANDON; ALLEMANG, 2020), por exemplo.

Em uma outra tentativa de ilustrar o conceito da retroalimentação de dados entre a base de dados de uma instituição cultural e o ecossistema wiki, denominado em inglês como "*roundtripping*", termo que será utilizado ao longo deste projeto, a pesquisadora Sandra Fauconnier (2019) produziu a ilustração apresentada na figura 02. A imagem destaca as ações esperadas da instituição cultural, ou seja, a publicação de novos itens, a inclusão de novos metadados nos itens já publicados e a inclusão de correções realizadas pela própria instituição nos metadados já publicados. De outra parte, a imagem também ressalta o que se espera de colaboração oriunda do Wikidata e do Wikimedia Commons para a base de dados da instituição, ou seja, as traduções, novos metadados sugeridos pelos usuários e correções realizadas nos metadados já existentes.

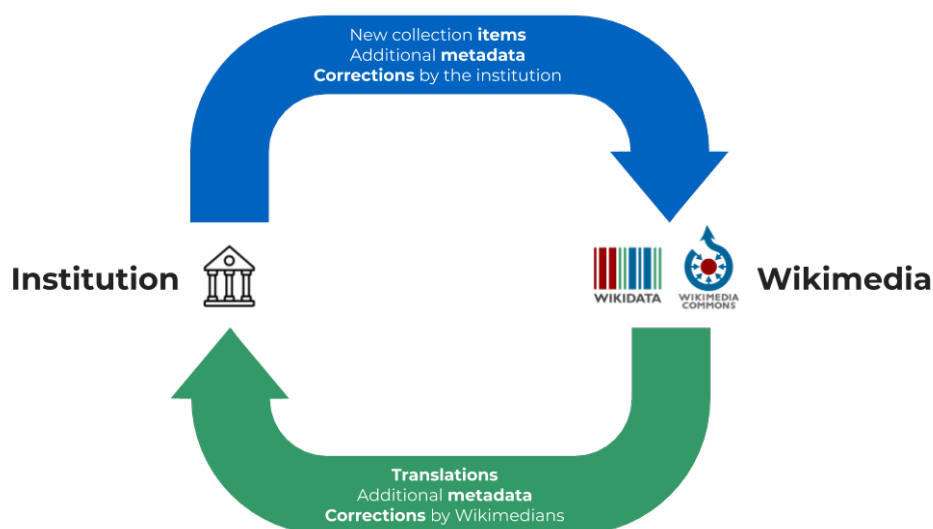


Figura 02. Fluxo de retroalimentação de dados entre base de dados de uma instituição cultural e a Wikimedia Commons e Wikidata. Fonte: (FAUCONNIER, 2019).

Cientes desse problema e do potencial desse tipo de serviço para as instituições culturais, a agência responsável pela política do patrimônio cultural do governo sueco, a *Swedish National Heritage Board* (2021), iniciou uma pesquisa sobre a implementação de um serviço para devolver às instituições de origem os metadados melhorados por usuários no ambiente do Wikimedia Commons. No relatório publicado com os primeiros resultados da pesquisa, Zeinstra (2019) relata que o principal interesse identificado nas instituições culturais por esse tipo de serviço seria a possibilidade de se beneficiarem das seguintes funcionalidades: receber novos tipos de metadados que possam ser incluídos em suas bases de dados, receber e avaliar a pertinência de informação sobre metadados alterados (correção sintática, identificação de erros), receber metadados traduzidos para outras línguas, receber e avaliar a sugestão de novas categorizações e classificações temáticas dos seus arquivos de mídia, receber e avaliar a alteração digital em suas mídias (melhoria de qualidade, contraste, ruído, entre outros). Ao final do relatório, o autor faz a recomendação de que, para viabilizar esse serviço, se faz necessário reduzir a barreira técnica de adoção a partir da criação de novas funcionalidades simples de exportação de dados das bases de dados das instituições e a adoção de APIs (*Application Program Interface*) de comunicação entre os sistemas e o Wikimedia Commons. Já no relatório final da pesquisa, Larsson, Ånas, e Zeinstra (2019) relatam que uma das maiores dificuldades encontradas nos estudos de caso realizados com as instituições suecas foi o fato de elas utilizarem sistemas de bases de dados diferentes, nos quais o processo de importação e exportação de dados tinha de ser feito de forma customizada para cada um deles. Tal procedimento aumentava a complexidade técnica de cada etapa, exigindo soluções personalizadas para cada base de dados, diminuindo assim a capacidade de escalabilidade das soluções encontradas. A adoção de um padrão, seja de um sistema de gerenciamento de coleções para instituições culturais seja de uma interface de comunicação de dados entre sistemas, foi sugerida pelos autores.

Ao refletir sobre o cenário brasileiro e a possibilidade de avanço na presença das coleções das instituições culturais brasileiras no ecossistema wiki, algumas condições de contorno devem ser levadas em consideração. Inicialmente, há um número ainda relativamente baixo de instituições que utilizam um software específico para a catalogação de acervos. Segundo a pesquisa TIC Cultura do Comitê Gestor da Internet (2019), apenas 31% dos museus, 36% das bibliotecas e 30% dos pontos de cultura no país utilizam algum tipo de software de catalogação de acervos. Melhor posicionados nesse ranking estão os arquivos, com 63% do total. Percebe-se que a maioria dos equipamentos ainda carecem de estratégias mais sistematizadas de gestão de dados de sua documentação a respeito dos objetos culturais



custodiados. Reconhecendo as fragilidades estruturais do setor cultural em relação à internet no Brasil, diversas foram as iniciativas e projetos que tentaram propor políticas e modelos de governança para a área dos acervos digitais culturais no país (DIAS e MARTINS, 2020). Dentre as várias iniciativas brasileiras, destaque-se para a presente proposta o projeto Tainacan.

O Tainacan é um software livre para a gestão e difusão de coleções digitais de instituições culturais na internet desenvolvido pela Universidade Federal de Goiás e pela Universidade de Brasília. Conta atualmente com fomento do Instituto Brasileiro de Museus (IBRAM), da Fundação Nacional das Artes (FUNARTE), do Instituto do Patrimônio Histórico e Artístico Nacional (IPHAN) e do Governo do Estado do Espírito Santo. O software é desenvolvido como um plugin para WordPress, permitindo tornar o sistema um repositório digital completo para a gestão e difusão de acervos. O Tainacan é cada vez mais utilizado para a catalogação, gestão e difusão de acervos digitais na internet de instituições culturais no Brasil. Um dos fatores que se destaca na sua crescente adoção é ser um software gratuito e desenvolvido no âmbito de várias iniciativas de políticas públicas culturais. Outro fator, não menos importante, é o fato de ser desenvolvido em uma tecnologia bastante disseminada no território brasileiro, o ambiente WordPress, muito utilizado no país para desenvolvimento de sites na web. Esse conhecimento gera uma curva de adoção mais reduzida, dada a maior quantidade de profissionais espalhados pelo país que conhecem e possuem experiência em customização de sites na plataforma WordPress.

Segundo as estatísticas de *download* do software, fornecidas pelo repositório de distribuição do WordPress (2021), o Tainacan já foi baixado mais de 9.000 vezes e possui mais de 400 instalações ativas. Dos 30 museus que são ligados oficialmente ao IBRAM, 20 deles já utilizam o Tainacan e disponibilizam mais de 15.000 itens de objetos culturais na internet (INSTITUTO BRASILEIRO DE MUSEUS, 2021). O software também vem suscitando o interesse de instituições universitárias para o uso na gestão de coleções. Martins e Martins (2020) apresentam um estudo no qual identificaram 17 instituições universitárias no Brasil, no México e nos Estados Unidos utilizando o Tainacan para a gestão de suas coleções, representando um total de 63 coleções com mais de 11.890 itens catalogados. Além disso, existem diversos relatos recentes na literatura científica da área da Ciência da Informação descrevendo estudos de caso de implantação e migração de outros sistemas de gestão de coleções digitais para o Tainacan. Martins *et al.* (2017) apresentam o caso da revista Filme e Cultura, importante revista da área do cinema brasileiro editada pelo Ministério da Cultura, cujos primeiros números remontam aos anos 1960 sendo disponibilizados na íntegra online com o Tainacan. Martins, Carmo e Germani (2018) também descrevem a migração do acervo

museológico do Museu do Índio, instituição com sede no Rio de Janeiro e ligada a FUNAI, em que foram migrados mais de 18.000 itens do software PHL para o Tainacan. Oliveira e Martins (2019) relatam o estudo de caso da migração do acervo da Fundação Nacional das Artes em que detalham as etapas da análise da informação do acervo existente, do tratamento dos dados e da implementação de 7 coleções com mais de 2.500 itens disponibilizados para acesso ao público com o Tainacan. O interesse acadêmico de pesquisa e implementação do Tainacan em diversos projetos científicos também pode ser percebido pelo número de documentos que mencionam o software na plataforma Google Acadêmico. No momento de produção deste projeto de pesquisa<sup>2</sup>, a plataforma já mencionava 151 documentos indexados, sendo que no ano de 2015 há apenas um documento e no ano de 2019, 45 documentos mencionando o projeto.

A partir do exposto, é possível perceber que o Tainacan vem se constituindo como um software amplamente utilizado por instituições culturais no Brasil, viabilizando-se como uma plataforma com uma larga base de acervos culturais já disponíveis na internet. Isso torna o Tainacan um software potencialmente viável para funcionar como um ponto de partida para a ampliação da presença dos acervos culturais brasileiros no ecossistema wiki. É a partir desse contexto que se evidencia o problema a ser trabalhado no presente projeto de pesquisa: **como implementar no software livre Tainacan as funcionalidades técnicas para a publicação de metadados de seus itens catalogados no Wikidata e das mídias representativas dos itens no Wikimedia Commons? Reciprocamente, como gerar indicadores e métricas para que edições realizadas nos projetos Wikimedia possam ser monitoradas e notificadas aos gestores das coleções por meio do próprio Tainacan e, se consideradas válidas, eventualmente integradas à base local?** Em síntese, deseja-se, com o presente projeto, implementar todas as funcionalidades para a realização do *roundtripping* entre as coleções digitais das instituições culturais disponibilizadas no Tainacan e o ecossistema wiki.

Metodologicamente, o projeto pretende realizar um estudo de caso a partir das funcionalidades de *roundtripping* que serão implementadas no software Tainacan, com o acervo museológico do Museu Paulista da Universidade de São Paulo. A equipe desse museu já possui um projeto de experimentação com o Tainacan desde o ano de 2020 e desde 2017 tem uma iniciativa ratificada de difusão digital nos projetos Wikimedia (PESCHANSKI, 2021). Em 2021, o Museu Paulista realizou um curso de formação da equipe na tecnologia e acompanhou

---

2 [https://scholar.google.com.br/scholar?q=%22Tainacan%22+&hl=pt-BR&as\\_sdt=0,5](https://scholar.google.com.br/scholar?q=%22Tainacan%22+&hl=pt-BR&as_sdt=0,5). Acesso em 09 de jun. 2021.

um ensaio com os pesquisadores da Universidade de Brasília da migração de sua documentação museológica para o Tainacan. Entende-se que um estudo de caso de implementação dos resultados do presente projeto junto ao Museu Paulista possa demonstrar os impactos da publicação dos acervos no ecossistema wiki e as possibilidades de monitoramento do seu uso e ações de colaboração realizadas pelos usuários. O Museu Paulista é especialmente apto ao estudo de caso aqui proposto. Trata-se de um museu de história universitário e público que tem investido de maneira consistente na documentação de suas vastas e diversificadas coleções (são mais de 70 mil imagens em suportes variados, 30 mil objetos e 200 metros lineares de documentação textual), o que resulta em metadados de origem bem estruturados. As suas coleções são objetos de pesquisas acadêmicas cujos resultados em dissertações, teses e artigos científicos têm sido mobilizados na elaboração de novos verbetes e melhoria de verbetes já existentes da Wikipedia, qualificando assim não só o acervo institucional, mas também a circulação e o reuso em redes de informação. Outro aspecto de interesse é o fato dos objetos, imagens e escritos preservados pelo museu serem oriundos de segmentos sociais distintos por atenderem a uma política de aquisição que ambiciona representatividade social, o que resulta em uma relação empática do público com esse tipo específico de patrimônio documental preservado. Por fim, é preciso destacar a missão educacional do Museu Paulista, que não se restringe ao nível superior, mas abrange também as redes de ensino fundamental e médio. Trata-se, portanto, de instituição que tem como visão a valorização do patrimônio material de modo integrado à sua missão educativa.

## 2. Resultados esperados

O **primeiro resultado esperado** é a especificação de um **modelo conceitual de como deve ocorrer o *roundtripping*** entre a base de dados da instituição cultural e o ecossistema wiki. Esse modelo deve presumir como devem ocorrer as ações de publicação dos metadados no Wikidata e das mídias no Wikimedia Commons. Essa tarefa deve levar em consideração a necessidade de adequação aos formatos técnicos da web semântica exigidos pelo ecossistema wiki, conforme mencionado na seção anterior, além da questão do controle de permissão para acesso e publicação de dados entre os dois lados do *roundtripping*. Outra tarefa a ser modelada é o monitoramento das edições feitas pelos usuários nos itens publicados e nos dados de uso, além da visualização dos itens. Deve-se especificar o que deve ser monitorado, como esse monitoramento se dará, sua periodicidade e formato de agregação e desagregação de dados de forma a favorecer a análise.

O **segundo resultado esperado** é o desenvolvimento de um **plugin para WordPress complementar ao Tainacan** que implemente as funcionalidades de publicação de metadados dos objetos catalogados no Tainacan no Wikidata, das mídias no Wikimedia Commons e permita monitorar as edições e contribuições realizadas nesses recursos de informação. Além disso, espera-se também monitorar os diferentes tipos de usos que serão feitos nessas mídias, tais como as visualizações mensais e os verbetes na Wikipédia em que foram utilizadas. É importante ressaltar que o plugin será complementar ao Tainacan, permitindo que a instituição cultural apenas o instale se tiver interesse de publicar seu acervo no ecossistema wiki e, caso contrário, utilize o Tainacan apenas para a gestão de seu acervo institucional. Essa funcionalidade também facilitará a adoção do novo plugin pelas instituições que já utilizam o Tainacan, que terão apenas que instalar um plugin complementar. Cabe ressaltar que tal plugin será também distribuído como software livre junto ao repositório de plugins do WordPress, tal como hoje acontece com o Tainacan.

O **terceiro resultado esperado** é a documentação e os indicadores de impacto do **estudo de caso a ser realizado junto ao Museu Paulista**. O estudo de caso será realizado em conjunto com a equipe do museu, o que presume uma etapa de capacitação, planejamento, execução e monitoramento dos resultados, bem como a coleta dos indicadores quantitativos de impacto e qualitativos da percepção da equipe técnica e de gestão do museu sobre os efeitos de tal serviço na realização de suas atividades. Espera-se que essa documentação possa servir de apoio e incentivo a adoção desse tipo de serviço por outras instituições culturais no Brasil.

### **3. Desafios científicos e tecnológicos e os meios e métodos para superá-los**

Identifica-se, para a realização do presente projeto, **4 desafios científicos e tecnológicos**. O **primeiro desafio** tem relação com a **tarefa de reconciliação dos valores dos metadados** dos itens a serem publicados. O **segundo desafio** refere-se ao **fluxo de publicação** de dados. O **terceiro desafio** diz respeito ao **fluxo de monitoramento das edições e colaborações** dos usuários nos itens publicados. O **quarto** desafio está relacionado aos **indicadores e métricas** que devem ser utilizados para monitorar o impacto dos resultados do **estudo de caso**. Os 4 desafios são detalhados a seguir.

No **primeiro desafio**, ressalta-se a questão da **reconciliação dos valores dos metadados** dos objetos catalogados no Tainacan com valores existentes no Wikidata. A reconciliação é a tarefa que permite reconhecer se um valor específico de um metadado é uma entidade já existente no Wikidata e atribuir a esse valor a identificação dessa entidade. Para ilustrar o desafio parte-se do seguinte exemplo: um metadado denominado "autor", criado para catalogar itens em uma

instalação do Tainacan, foi preenchido com o valor "Machado de Assis" em um determinado item. Ao enviá-lo para publicação no Wikidata, a atividade de reconciliação deve reconhecer que esse valor é uma entidade já existente na base, no caso a entidade Q311145 (<https://www.wikidata.org/wiki/Q311145>), e publicar o item já ligando seu valor à entidade reconhecida. Essa tarefa permite um controle semântico dos dados publicados, aplicando as boas práticas de dados abertos ligados da web semântica na interligação e anotação de dados sob licença aberta por meio de um Uniform Resource Identifier (URI). Desse modo, tornar-se-ia possível o controle de autoridades e entidades semânticas de uma forma geral. Para superar tal desafio, será necessário implementar, no plugin a ser desenvolvido, uma funcionalidade que permita identificar entidades no Wikidata que, ou sejam idênticas textualmente ou sejam aproximadas do valor dos metadados a serem reconciliados no Tainacan. Tal tarefa deverá prever uma instância de decisão humana de aprovação da reconciliação, dado que há um potencial para ambiguidades e homônimos em valores de texto. Para reduzir tal problema, a abordagem que vem sendo utilizada por outros softwares que implementam serviços similares de reconciliação, tais como o Open Refine<sup>3</sup>, identifica inicialmente o tipo do metadado a ser reconciliado, ou seja, reconhece seu significado semântico atribuído explicitamente pelo gestor do acervo. Por exemplo, sabendo que um metadado é do tipo "autor" na Wikidata, os valores a serem reconciliados reduzem de forma significativa sua abrangência. Uma vez reconhecendo o tipo de metadado a ser reconciliado, deve-se calcular a proximidade entre os valores dos metadados no Tainacan com as entidades daquele tipo na Wikidata. Para isso, algoritmos como o FuzzyWuzzy<sup>4</sup>, também utilizado pelo software Open Refine para seu serviço de reconciliação, podem ser utilizados. O FuzzyWuzzy calcula a distância de Levenshtein (NAVARRO, 2001) entre dois conjuntos de texto e apresenta a probabilidade, variando de 0 a 100, de sua proximidade.

No **segundo desafio**, destaca-se o **fluxo de publicação de dados**. Para o presente projeto, deve-se pressupor uma forma de comunicação para publicação dos metadados produzidos no Tainacan no Wikidata, e das mídias anexadas aos itens catalogados no Tainacan na Wikimedia Commons. O Wikidata<sup>5</sup> oferece quatro formas de publicação de dados em sua base: manual item a item na interface do sistema, usando ferramentas online tais como Mix'n'match e Quickstatements 2, por meio de uma API ou por meio do desenvolvimento de um robô para a publicação automática de dados. Entende-se que, uma vez os valores dos metadados tenham

---

<sup>3</sup> <https://wikidata.reconci.link/>. Acesso em 09 jun. 2021.

<sup>4</sup> <https://github.com/seatgeek/fuzzywuzzy> Acesso em 09 jun. 2021.

<sup>5</sup> [https://www.wikidata.org/wiki/Wikidata:Data\\_donation](https://www.wikidata.org/wiki/Wikidata:Data_donation). Acesso em 09 jun. 2021.

sido reconciliados, conforme descritos no primeiro desafio, e o tipo de metadados identificado a partir de seu significado semântico pelos gestores da coleção, a publicação de metadados no Wikidata poderia ser realizada por meio de um robô a ser implementado no plugin proposto. No Wikimedia Commons<sup>6</sup>, também é possível a construção de um robô para a publicação de dados de forma automática. Espera-se a implementação desse robô junto ao plugin proposto. Uma vez que os dados tenham sido reconciliados e publicados, deve-se guardar seus identificadores únicos (URI semântico) no banco de dados do Tainacan como mais um metadado de cada item. Será a partir desse registro que o monitoramento das edições e das visualizações poderá ser realizado.

No **terceiro desafio**, salienta-se o **fluxo de monitoramento das edições e colaborações** realizadas nos itens publicados no Wikidata e no Wikimedia Commons. Os dois serviços implementam um *feed* em formato ATOM/XML para o histórico de edições e colaborações realizadas em um item. De forma a exemplificar seu funcionamento e como ele pode ser integrado ao Tainacan, apresenta-se a seguir a forma como o histórico de edições e colaborações pode ser visualizado no caso de um objeto já publicado do Museu Paulista. Trata-se do objeto denominado "Porta-retrato"<sup>7</sup>, publicado no Wikimedia Commons e que retrata Santos Dumont. Junto a esse objeto, pode-se encontrar o seu histórico<sup>8</sup> de edições. Esse histórico também pode ser acessado para consumo computacional por meio do *feed*<sup>9</sup> em formato ATOM/XML. Os metadados descritivos do objeto se encontram na Wikidata<sup>10</sup> e, da mesma forma, é possível visualizar seu histórico<sup>11</sup> e ter acesso ao formato para consumo computacional por ATOM/XML<sup>12</sup>. Logo, entende-se que o monitoramento das edições e colaborações realizadas possa ser feito por meio da leitura assíncrona desses dados e pela exibição dos mesmos em um painel de dados que possa ser visualizado e avaliado pelos gestores da instituição cultural. A incorporação ou não das colaborações realizadas dependerá de uma avaliação humana, sendo que os dados dessas edições serão incorporados ou não a documentação dos itens conforme o resultado dessa avaliação.

---

<sup>6</sup> [https://commons.wikimedia.org/wiki/Commons:Guide\\_to\\_batch\\_uploading](https://commons.wikimedia.org/wiki/Commons:Guide_to_batch_uploading). Acesso em 09 jun. 2021.

<sup>7</sup> [https://commons.wikimedia.org/wiki/File:Porta-Retrato,\\_Acervo\\_do\\_Museu\\_Paulista\\_da\\_USP\\_\(18\).jpg](https://commons.wikimedia.org/wiki/File:Porta-Retrato,_Acervo_do_Museu_Paulista_da_USP_(18).jpg). Acesso em 09 jun. 2021.

<sup>8</sup> [https://commons.wikimedia.org/w/index.php?title=File:Porta-Retrato,\\_Acervo\\_do\\_Museu\\_Paulista\\_da\\_USP\\_\(18\).jpg&action=history](https://commons.wikimedia.org/w/index.php?title=File:Porta-Retrato,_Acervo_do_Museu_Paulista_da_USP_(18).jpg&action=history). Acesso em 09 jun. 2021.

<sup>9</sup> [https://commons.wikimedia.org/w/index.php?title=File:Porta-Retrato,\\_Acervo\\_do\\_Museu\\_Paulista\\_da\\_USP\\_\(18\).jpg&feed=atom&action=history](https://commons.wikimedia.org/w/index.php?title=File:Porta-Retrato,_Acervo_do_Museu_Paulista_da_USP_(18).jpg&feed=atom&action=history). Acesso em 09 jun. 2021.

<sup>10</sup> <https://www.wikidata.org/wiki/Q61766152>. Acesso em 09 jun. 2021.

<sup>11</sup> <https://www.wikidata.org/w/index.php?title=Q61766152&action=history>. Acesso em 09 jun. 2021.

<sup>12</sup> <https://www.wikidata.org/w/index.php?title=Q61766152&feed=atom&action=history>. Acesso em 09 jun. 2021.

Por fim, no **quarto desafio**, ressaltam-se os **indicadores e métricas de monitoramento** de uso e visualização das mídias em verbetes e páginas da Wikipédia. A Wikimedia Foundation disponibiliza um serviço de dados<sup>13</sup> para coleta e monitoramento de estatísticas dividido em três grandes conjuntos de métricas: leitura, contribuições e conteúdo. Essas métricas podem ser acessadas para consumo computacional por meio de um serviço denominado *Analytics Query Service* (AQS)<sup>14</sup> que oferece um conjunto de APIs. Por exemplo, existe uma API dedicada apenas ao monitoramento de visualizações de página, denominada Pageview<sup>15</sup> API. Pretende-se utilizar esses serviços para monitorar os objetos publicados no Wikimedia Commons e Wikidata e como eles foram reutilizados para ilustrar páginas e verbetes nas Wikipédias de diferentes idiomas.

#### 4. Cronograma

O projeto será desenvolvido ao longo de 24 meses, conforme as etapas descritas a seguir, que acontecerão consecutivamente:

##### a. Fase 1 - duração 4 meses

**Revisão sistemática de literatura científica e relatórios técnicos sobre o conceito e aplicações de *roundtripping*.** Essa fase consiste na elaboração de um protocolo de revisão sistemática de literatura científica em bases de publicações indexadas nacionais e internacionais com revisão por pares e na análise de relatórios técnicos da Wikimedia Foundation e outros atores que vêm desenvolvendo iniciativas nessa direção. O objetivo é estabelecer um estado da arte para coletar insumos que possam impactar no desenvolvimento técnico do projeto.

##### b. Fase 2 – duração 4 meses

**Modelagem conceitual do serviço de *roundtripping* entre o software livre Tainacan, o Wikidata e o Wikimedia Commons.** Essa fase consiste na especificação conceitual e operacional de como o serviço deverá funcionar. Tem por objetivo desenhar tecnicamente as funcionalidades que devem ser contempladas no Tainacan, o que deve estar na interface de comunicação entre os sistemas e o que deve ser realizado no Wikidata e no Wikimedia Commons. A fase realizará um desenho técnico operacional do serviço que será a base para a implementação do plugin em fase posterior. Vale ressaltar que para essa fase, os técnicos e os gestores do Museu Paulista serão ouvidos para discutir ideias e funcionalidades que

---

<sup>13</sup> <https://stats.wikimedia.org/#/metrics/all-projects>. Acesso em 09 jun. 2021.

<sup>14</sup> <https://wikitech.wikimedia.org/wiki/Analytics/Systems/AQS>. Acesso em 09 jun. 2021.

<sup>15</sup> <https://wikitech.wikimedia.org/wiki/Analytics/AQS/Pageviews>. Acesso em 09 jun. 2021.

poderiam ser potencialmente úteis ao seu trabalho cotidiano de gestão da informação da documentação museológica.

**c. Fase 3 - duração 6 meses**

**Codificação do plugin para WordPress.** Nessa fase será implementado o plugin conforme o modelo conceitual especificado na fase anterior.

**d. Fase 4 - duração 6 meses**

**Estudo de caso com o Museu Paulista.** Nessa fase será realizado o estudo de caso junto ao Museu Paulista. A fase contempla a realização de uma etapa de capacitação dos técnicos e gestores do museu nas funcionalidades do plugin. Também envolve a escolha de um conjunto de dados a ser publicado usando os recursos técnicos desenvolvidos. Por fim, será feita a coleta e o monitoramento de métricas e indicadores de impacto para se avaliar os efeitos do uso do plugin. Também inclui etapa qualitativa de entrevista com os participantes do caso de uso.

**e. Fase 5 - duração 4 meses**

**Seminário de divulgação científica, produção de relatórios e artigos científicos do projeto:** consiste na documentação sistematizada em relatórios técnicos e artigos científicos relatando os principais resultados e descobertas realizadas no projeto. Nessa etapa, também será realizado um seminário de divulgação científica dos resultados alcançados e aprendizados obtidos ao longo do projeto no Museu Paulista, em São Paulo.

## **5. Disseminação e avaliação**

Espera-se que os resultados do projeto sejam disseminados a partir dos seguintes produtos:

- **Software livre - plugin para WordPress – roundtripping Tainacan – Wikidata – Wikimedia Commons:** será disponibilizado um plugin no repositório de plugins do WordPress para download gratuito e uso por outras instituições culturais do Brasil e exterior. Pode-se monitorar o número de downloads e de instalações ativas utilizando o plugin ao longo do tempo como uma das formas de se avaliar o resultado de adesão e interesse na tecnologia.
- **Modelo conceitual do *roundtripping*:** será publicado em relatório técnico do modelo conceitual do serviço, permitindo a socialização, a discussão e a revisão por pares e interessados em tal desenvolvimento tecnológico. Entende-se que o modelo conceitual



favoreça o avanço da compreensão de como essas redes devem interoperar e as funcionalidades de cada uma das partes.

- **Artigos científicos:** serão desenvolvidos e submetidos para publicação ao menos 4 artigos científicos para publicação preferencial em periódicos internacionais da área da Ciência da Informação e áreas correlatas. Espera-se um artigo com os resultados da revisão sistemática de literatura, um artigo do modelo conceitual, um artigo descritivo do plugin, suas funcionalidades e possibilidades operacionais e um artigo sobre os resultados do estudo de caso.
- **Seminário de divulgação dos resultados:** o projeto pretende realizar um seminário para a socialização dos resultados. Pretende-se a montagem de duas mesas para divulgação dos resultados conceituais e teóricos do projeto. Também se pretende a realização de uma oficina aberta à inscrição de participantes para experimentação da tecnologia por meio de um tutorial dirigido.

## 6. Outros apoios

O Instituto Brasileiro de Museus por meio de termo aditivo<sup>16</sup> no valor de R\$500.000,00 realizado no final do ano de 2020 e com vigência até o final do ano de 2021 se compromete apoiar diversas iniciativas de continuidade de desenvolvimento do projeto Tainacan, constando como uma das ações a disponibilização do padrão de metadados INBCM (Inventário Nacional de Bens Culturais Musealizados)<sup>17</sup> e o Tesouro de Objetos do Patrimônio Cultural nos Museus Brasileiros<sup>18</sup> no Wikidata e Wikimedia Commons. Ressalta-se que tais recursos são complementares ao projeto e podem beneficiar os esforços aqui propostos.

O Museu Paulista da Universidade de São Paulo se compromete por meio de carta, anexa ao projeto, a ceder técnicos e gestores para participação no projeto nas etapas de definição do modelo conceitual e estudo de caso. Além disso, o Museu se compromete a ceder espaço e operacionalizar a logística para realização do seminário de divulgação dos resultados do projeto. A Associação Wikimedia no Brasil formalizou interesse na proposta, em carta anexa.

---

<sup>16</sup> [https://www.museus.gov.br/wp-content/uploads/2021/01/SEI\\_IBRAM-1120938-7-Termo-Aditivo-TED-Tainacan.pdf](https://www.museus.gov.br/wp-content/uploads/2021/01/SEI_IBRAM-1120938-7-Termo-Aditivo-TED-Tainacan.pdf). Acesso em 09 jun. 2021.

<sup>17</sup> <https://www.museus.gov.br/resolucao-regulamenta-inventario-nacional-de-bens-culturais-musealizados/>. Acesso em 09. Jun 2021.

<sup>18</sup> <http://www.tesauromuseus.com.br/>. Acesso em 09 jun. 2021.

## 7. Bibliografia

ASSOCIATION OF RESEARCH LIBRARIES. **ARL White Paper on Wikidata Opportunities and Recommendations**. ARL Task Force on Wikimedia and Linked Open Data, Washington: ARL, 2019. Disponível em: <https://www.arl.org/wp-content/uploads/2019/04/2019.04.18-ARL-white-paper-on-Wikidata.pdf>. Acesso em: 08 de jun. 2021.

CARMO, Danielle; MARTINS, Dalton Lopes. Os Acervos Culturais Brasileiros no Repositório Wikimedia Commons. In: WIDAT, 2019, Brasília. **Anais eletrônicos...** Brasília: UnB, 2019a. p.121- 127. Disponível em: <http://widat2019.fci.unb.br/index.php/anais-widat-2019>. Acesso em 20 mai. 2021.

CARMO, Danielle; MARTINS, Dalton Lopes. A presença dos museus brasileiros na ecologia informacional da Fundação Wikimedia: estudo de caso do projeto Sum of All Paintings. In: ENANCIB, 20., 2019, Florianópolis. **Anais eletrônicos...**, Florianópolis: UFSC, 2019b, p.1-20. Disponível em: <https://conferencias.ufsc.br/index.php/enancib/2019/paper/view/909/679>. Data de acesso: 18 mai. 2021.

COMITÊ GESTOR DA INTERNET NO BRASIL. **TIC Cultura 2018: Pesquisa sobre o uso das Tecnologias de Informação e Comunicação nos equipamentos culturais brasileiros**, São Paulo: Comitê Gestor da Internet no Brasil, 2019. Disponível em: [https://cetic.br/media/docs/publicacoes/1/tic\\_cultura\\_2018\\_livro\\_eletronico.pdf](https://cetic.br/media/docs/publicacoes/1/tic_cultura_2018_livro_eletronico.pdf). Acesso em: 19 maio. 2021.

DIAS, Calíope Victor Spíndola de Miranda; MARTINS, Dalton Lopes. Iniciativas brasileiras em torno da construção de uma política nacional para acervos digitais de instituições de memória: o desafio da memória em tempos de cultura digital. **Políticas Culturais em Revista**, Salvador, v. 13, n. 1, p. 16-46, jan./jun. 2020. Disponível em: <https://portalseer.ufba.br/index.php/pculturais/article/view/35616/21211>. Acesso em: 09 jun. 2021.

FAUCONNIER, Sandra. **Data Roundtripping: a new frontier for GLAM-Wiki collaborations**. 2019. Disponível em: <https://diff.wikimedia.org/2019/12/13/data-roundtripping-a-new-frontier-for-glam-wiki-collaborations/>. Acesso em: 20 maio 2021.

FERRITER, Meghan. **Integrating Wikidata at the Library of Congress**. The Signal: Blog of The Library of Congress. 2009. Disponível em: <https://blogs.loc.gov/thesignal/2019/05/integrating-wikidata-at-the-library-of-congress/>. Acesso em: 08 de jun. 2021.

HENDLER, James; GANDON, Fabien; ALLEMANG, Dean. **Semantic Web for the Working Ontologist: Effective Modeling for Linked Data, RDFS, and OWL**. Morgan & Claypool, 2020.

INSTITUTO BRASILEIRO DE MUSEUS. **Acervo em rede**. 2021. Disponível em: <https://www.museus.gov.br/acesoainformacao/acoes-e-programas/acervo-em-rede/>. Acesso em: 08 de jun. 2021.

LARSSON, Albin; ÅNÄS, Susanna; ZEINSTRA, Maarten. **ARL Wikimedia Commons Data Roundtripping- Final Report**. 2019. Disponível em:

[https://upload.wikimedia.org/wikipedia/commons/e/e8/Wikimedia\\_Commons\\_Data\\_Roundtripping\\_-\\_Final\\_report.pdf](https://upload.wikimedia.org/wikipedia/commons/e/e8/Wikimedia_Commons_Data_Roundtripping_-_Final_report.pdf). Acesso em: 20 maio 2021.

MARTINS, Dalton Lopes; CARMO, Danielle do. Dinâmica de participação social na construção coletiva de informação no campo museal: estudo de caso dos museus na Wikipédia no âmbito do Instituto Brasileiro de Museus. **Liinc em Revista**, v. 15, n. 1, p. 140-159, jun. 2019. Disponível em: <http://revista.ibict.br/liinc/article/view/4607/4143>. Acesso em: 20 mai. 2021.

MARTINS, Dalton Lopes; CARMO, Danielle do; GERMANI, Leonardo Barbosa. Museu do índio: estudo de caso do processo de migração e abertura dos dados ligados semânticos do acervo museológico com o software livre tainacan. **Informação & Tecnologia**, v. 5, n. 2, p. 142-162, 2018. Disponível em: <https://periodicos.ufpb.br/index.php/itec/article/view/44590/27749>. Acesso em: 16 jun. 2021.

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. 30esp, p. 34-61, 2020. Disponível em: <https://www.revistas.usp.br/cpc/article/view/173022>. Acesso em: 9 jun. 2021.

MARTINS, Dalton Lopes; SEGUNDO, José Eduardo Santarém; SILVA, Marcel Ferrante; SIQUEIRA, Joyce. Repositório digital com o software livre tainacan: revisão da ferramenta e exemplo de implantação na área cultural com a revista filme cultura. ENANCIB, 18., 2017, Marília. **Anais eletrônicos...**, Marília: UNESP, 2017. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/105154>. Acesso em: 16 jun. 2021.

MONTEIRO, Juliana. Wikipédia e museus: uma parceria possível? **Pragmatizes - Revista Latino Americana de Estudos em Cultura**, Rio de Janeiro, n. 16, p. 62-73, 2019. Disponível em: <https://periodicos.uff.br/pragmatizes/article/viewFile/27509/16802>. Acesso em: 20 maio 2021.

NAVARRETE, Trilce; VILLAESPESA, Elena. Image-based information: paintings in Wikipedia. **Journal of Documentation**. v.77, n.2, p. 359-380, 2021. Disponível em: [https://www.emerald.com/insight/0022-0418](https://www.emerald.com/insight/0022-0418.htm).htm. Acesso em: 21 mai. 2021.

NAVARRO, Gonzalo. A guided tour to approximate string matching. **ACM Computing Surveys**. v.33, n.1, p.31-88. 2001. Disponível em: <https://dl.acm.org/doi/abs/10.1145/375360.375365>. Acesso em: 21 mai. 2021.

OLIVEIRA, Luís Felipe Rosa de; Martins, Dalton Lopes. Fundação Nacional de Artes: Estudo de Caso da Migração e Publicação dos Dados do Acervo Digital com o Software Livre Tainacan. In: ENANCIB, 20., 2019, Florianópolis. **Anais eletrônicos...**, Florianópolis: UFSC, 2019, p.1-20. Disponível em: <https://conferencias.ufsc.br/index.php/enancib/2019/paper/view/909679>. Data de acesso: 18 mai. 2021.

OLIVEIRA, Luís Felipe Rosa de; Martins, Dalton Lopes. Explorando a Reconciliação de Dados Culturais na Wikidata: experimento aplicado com o acervo museológico do Museu Histórico Nacional. **Ciência da Informação**, [S. l.], v. 49, n. 3, 2020. Disponível em: <http://revista.ibict.br/ciinf/article/view/5462>. Acesso em: 22 jun. 2021.

PESCHANSKI, João Alexandre. Variedades de processos de difusão digital colaborativa: descrição e análise de iniciativas GLAM-Wiki no Brasil. **Resgate: Revista Interdisciplinar de Cultura**, Campinas, SP, v. 29, n. 1, p 1-28. 2021. Disponível em: <https://periodicos.sbu.unicamp.br/ojs/index.php/resgate/article/view/8659966>. Acesso em: 11 jun. 2021.

SWEDISH NATIONAL HERITAGE BOARD. Wikimedia Commons Data Roundtripping. Disponível em: [https://meta.wikimedia.org/wiki/Wikimedia\\_Commons\\_Data\\_Roundtripping](https://meta.wikimedia.org/wiki/Wikimedia_Commons_Data_Roundtripping). Acesso em: 18 mai. 2021.

VILLAESPESA, Elena; NAVARRETE, Trilce. Museum Collections on Wikipedia: Opening Up to Open Data Initiatives. **Museums and the Web 2019**. Boston, 2019. Disponível em: <https://mw19.mwconf.org/paper/museum-collections-on-wikipedia-opening-up-to-open-data-initiatives/>. Acesso em: 18 mai. 2021.

WIKIDATA. Página principal. Disponível em: <https://www.wikidata.org/> Acesso em: 08 de jun. 2021.

WIKIMEDIA COMMONS. **Página principal**. Disponível em: [https://commons.wikimedia.org/wiki/Main\\_Page](https://commons.wikimedia.org/wiki/Main_Page). Acesso em: 08 de jun. 2021.

WIKIMEDIA FOUNDATION. **Estatísticas da Wikimedia**. Disponível em: <https://stats.wikimedia.org/#/pt.wikipedia.org> . Acesso em: 08 de jun. 2021a.

WIKIMEDIA FOUNDATION. **List of Wikipedias**. Disponível em: [https://meta.wikimedia.org/wiki/List\\_of\\_Wikipedias](https://meta.wikimedia.org/wiki/List_of_Wikipedias). Acesso em: 08 de jun. 2021b.

WORDPRESS. **Estatísticas do Tainacan**. 2021. Disponível em: <https://br.wordpress.org/plugins/tainacan/advanced/>. Acesso em 09 de jun. 2021.

ZEINSTRA, Maarten. **Research Report - Returning commons community metadata additions and corrections to source**. Swedish National Heritage Board, 2019. Disponível em: [https://upload.wikimedia.org/wikipedia/commons/9/91/Research\\_Report\\_%E2%80%93\\_Returning\\_commons\\_community\\_metadata\\_additions\\_and\\_corrections\\_to\\_source.pdf](https://upload.wikimedia.org/wikipedia/commons/9/91/Research_Report_%E2%80%93_Returning_commons_community_metadata_additions_and_corrections_to_source.pdf). Acesso em: 20 maio 2021.