



Criação de um repositório temático no contexto das humanidades digitais para gestão da informação

Marcia Teixeira Cavalcanti¹, Ricardo Medeiros Pimenta², Josir Cardoso Gomes³

¹ Universidade Santa Úrsula (USU)/Faculdades Integradas Maria Thereza (FAMATH).
Email: marciacavalcanti@gmail.com

²Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT).
Email: ricardo.pimenta@ibict.br

³Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT).
Email: josircg@gmail.com

Resumo

Em 2017 foi fundado o Laboratório em Rede de Humanidades Digitais (Larhud), fruto do grupo de pesquisa que se dedica aos estudos críticos e às práticas em Humanidades Digitais. Sua criação teve como objetivo construir um canal estruturado de colaboração entre os pesquisadores internos e externos com foco no desenvolvimento e na reflexão sobre o pensar em soluções metodológicas para a pesquisa nas humanidades. Um caminho adotado pelo Larhud é criar um repositório digital temático, para reunir o conjunto da produção científica e acadêmica dos integrantes do GP, contribuindo para ampliar a visibilidade de sua produção e preservar sua memória. Um repositório pode ser definido como um local de armazenamento, de guarda e arquivamento de objetos digitais. Quanto ao seu propósito, se classifica como institucional – um espaço onde se armazenam os documentos administrativos, acadêmicos e científicos de uma instituição – ou temático – guarda objetos relacionados a uma determinada área sem limites institucionais – cada um com características e estrutura próprias. O artigo propõe apresentar um estudo descritivo, detalhando o passo a passo relacionado à estruturação das etapas de navegação, organização e representação da informação no contexto do repositório pela perspectiva da arquitetura da informação. A etapa inicial foi identificar seus usuários: administradores, pesquisadores e possíveis visitantes. A etapa atual consiste na organização e representação da informação a partir da definição das áreas temáticas, comunidades e coleções que irão compor estruturalmente o repositório, bem como da seleção do tipo de metadados que serão utilizados.

Palavras-chave: Repositório Digital, Humanidades Digitais, Arquitetura da Informação, Preservação Digital.



Abstract

In 2017 was founded the Laboratório em Rede de Humanidades Digitais (Larhud), the result of a research group dedicated to critical studies and practices in Digital Humanities. Its purpose was to build a structured channel of collaboration between internal and external researchers focused on the development and reflection on thinking about methodological solutions for research in the humanities. One way Larhud adopts is to create a thematic digital repository to bring together the scientific and academic production of the members of the GP, contributing to increase the visibility of its production and preserve its memory. A repository can be defined as a location for storing, storing and archiving digital objects. As for its purpose, it is classified as institutional - a space where administrative, academic and scientific documents of an institution - or thematic - store objects related to a certain area without institutional limits - each with its own characteristics and structure. The article proposes to present a descriptive study, detailing the step by step related to the structuring of the steps of navigation, organization and representation of the information in the context of the repository from the information architecture perspective. The initial step was to identify its users: administrators, researchers and potential visitors. The current stage consists of the organization and representation of information from the definition of thematic areas, communities and collections that will structure the repository, as well as the selection of the type of metadata that will be used.

Keywords: Digital Repository, Digital Humanities, Information Architecture, Digital Preservation.

Introdução

Segundo Dacos (2011), as humanidades digitais se referem ao conjunto das ciências humanas e sociais, às artes e às letras e também designam uma transdisciplina, portadora dos métodos, dos dispositivos e das perspectivas heurísticas ligadas ao digital no domínio das ciências humanas e sociais (DACOS, 2011). No Brasil, ainda que a busca pelo tema das Humanidades Digitais se mostre recorrente nos últimos anos, com o discurso da interdisciplinaridade aparecendo em boa parte das iniciativas apresentadas publicamente, é possível percebermos na prática que esta interdisciplinaridade não ocorre de forma efetiva quando o assunto é acesso à informação e reuso de dados gerados por ferramentas e métodos que se mostram

interdisciplinares por natureza.

A virada digital (*digital turn*) abalou, de certa forma, as estruturas do cenário acadêmico, levando professores e pesquisadores a se questionarem e reavaliarem suas estratégias de preservação e de uso da informação científica para a continuidade da produção do conhecimento e da promoção ao seu acesso.

Dentro desse novo cenário que se apresentou foi criado, no ano de 2017, o Laboratório em Rede de Humanidades Digitais (Larhud), fruto do grupo de pesquisa que se dedica aos estudos críticos e às práticas em Humanidades Digitais, mas o cenário também nos convidou a pensar sobre a construção de um repositório que fosse capaz de guardar e preservar dados brutos (*raw data*) provenientes da pesquisa em humanidades de maneira a promover seu acesso pela comunidade científica e que possibilite, entre outras coisas, seu acompanhamento, a troca de informações além de, conseqüentemente, a geração de inovação e de novos conteúdos. Assim, o Larhud se propôs criar um repositório digital que tem por objetivo reunir o conjunto da produção científica e acadêmica dos integrantes do GP, contribuindo para ampliar a visibilidade de sua produção e também preservar sua memória.

O presente artigo pretende apresentar um estudo de caso descritivo, desenvolvido no âmbito do GP, em que será detalhado o passo a passo relacionado à estruturação das etapas de navegação, organização e representação da informação no contexto do repositório pela perspectiva da Arquitetura da Informação. A etapa inicial foi organizar uma reunião para responder a algumas questões básicas que dariam o suporte para se pensar a estrutura do repositório. O passo seguinte foi identificar seus usuários: administradores, pesquisadores e possíveis visitantes, definindo para cada um deles um nível de permissões. A etapa atual consiste na organização e representação da informação a partir da definição das áreas temáticas, comunidades e coleções que irão compor estruturalmente o repositório, bem como da seleção do tipo de metadados que serão utilizados.

O início de tudo

Um repositório pode ser definido, no contexto atual, como um local de armazenamento, de guarda e arquivamento de objetos digitais. Os repositórios, quanto ao seu propósito, são classificados como institucionais – um espaço onde se armazenam os documentos administrativos, acadêmicos e científicos de uma instituição – ou temáticos – guarda objetos relacionados a uma determinada área sem limites institucionais, como proposto pelo Larhud para a área das Humanidades Digitais – cada um com características e estrutura próprias.

Os repositórios digitais (RDs) são bases de dados online que reúnem de



maneira organizada a produção científica de uma instituição ou área temática. Os RDs armazenam arquivos de diversos formatos. Ainda, resultam em uma série de benefícios tanto para os pesquisadores quanto às instituições ou sociedades científicas, proporcionam maior visibilidade aos resultados de pesquisas e possibilitam a preservação da memória científica de sua instituição. Os RDs podem ser institucionais ou temáticos. Os repositórios institucionais lidam com a produção científica de uma determinada instituição. Os repositórios temáticos com a produção científica de uma determinada área, sem limites institucionais. (IBICT, s.d.)

A ideia de criação de um repositório temático para o Larhud surgiu após os pesquisadores perceberem que estavam gerando um número significativo de dados em suas pesquisas, mas não tinham um local em que estes dados pudessem ser preservados e compartilhados. Mas também ficou perceptível que os dados, enquanto fontes de pesquisa em potencial, não estavam recebendo o tratamento ideal, correndo mesmo o risco de se tornarem irrecuperáveis.

Do "crítico cultural" ao "processador de dados", é fato que este novo contexto de mudança vivida pelo pesquisador das humanidades lato sensu representa um novo "despertar" para um desafio razoavelmente difícil que se baseia no desenvolvimento de competências informacionais aplicadas às novas tecnologias de informação e comunicação (NTIC) por pesquisadores das humanidades. (PIMENTA, 2017, p.12)

Uma das indagações iniciais por parte dos pesquisadores envolvidos na criação do repositório foi sobre a plataforma a ser usada. A dúvida inicial era sobre o uso do Dspace, um software para o gerenciamento de acervo digital, utilizado para implementação de repositórios institucionais, ou o Zenodo, um repositório de dados de pesquisa criado pelo OpenAIRE e pelo *European Organization for Nuclear Research* (CERN) para fornecer um local confiável para os pesquisadores depositarem seus conjuntos de dados. Após algumas reuniões o Zenodo foi escolhido, e a decisão teve como parâmetros norteadores dois elementos principais: (1) disponibilidade de ser um repositório em «nuvem» em uma instituição de pesquisa confiável e não de mercado; (2) ser uma tecnologia de acesso aberto, além de gerar um identificador de objeto digital (DOI) que torna os dados depositados facilmente citáveis.

Posteriormente, considerando o resultado de reuniões ocorridas entre representantes do LARHUD e demais pesquisadores do IBICT Brasília, tomamos conhecimento do posicionamento daquela instituição em utilizar o Invenio – um software de código aberto usado para criação de repositórios, sendo um deles o próprio Zenodo. O apoio do IBICT foi fundamental para decidirmos pelo uso do Invenio na criação de um repositório «desenhado» de acordo com nossas escolhas e que ainda possui as

características do Zenodo descritas acima.

É preciso esclarecer que esta ainda é uma atividade em andamento dentro do Grupo de Pesquisa e, nesse escopo, não somente o «desenho» de como se dará tal repositório como sua implementação estão a ser desenvolvidos com a colaboração de todos os pesquisadores participantes da pesquisa.

Arquitetura da Informação

Vivemos na sociedade da informação, que significa dizer, de forma bem genérica, que vivemos em uma sociedade inaugurada pela internet e caracterizada por constantes mudanças ocasionadas pela tecnologia, o que passa a exigir de nós a habilidade de transformar o mar de informações que recebemos diariamente em conhecimento. Informações estas que nos chegam, ou deveriam chegar, de forma democratizada e universal e que são acessadas por nós via diferentes suportes tecnológicos e espaços informacionais digitais.

Quando pensamos no contexto acadêmico, as informações e os dados que os cientistas buscam encontram-se em sites, plataformas e repositórios, para citar apenas alguns destes espaços. Mas para que estas informações e dados possam ser localizados e recuperados para serem usados, é preciso que estes espaços informacionais tornem essas ações possíveis, ou seja, possuam “[...] uma arquitetura capaz de otimizar ou dificultar a experiência informacional dos sujeitos nestes ambientes” (GUSMÃO, 2018, p. 4).

A Arquitetura da Informação (AI) é a disciplina, ou deveria ser, que nos auxilia para a construção de espaços informacionais pelo qual são praticadas as ações de sistematizar e organizar a informação para que o usuário possa recuperá-la. Segundo Macedo (2005), o termo foi utilizado pela primeira vez pelo arquiteto Richard Saul Wurman em 1976, que o definia como a ciência e a arte de criar instruções para espaços organizados.

Na visão de Wurman a reunião, a organização e a apresentação da informação serviam a propósitos característicos aos das tarefas da Arquitetura. A Arquitetura da Informação seria uma expansão da profissão da Arquitetura, porém aplicada a espaços de informação. E as estruturas de informação influenciariam interações no mundo da mesma forma que as estruturas dos edifícios estimulam ou limitam as interações sociais. (MACEDO, 2005, p. 105)

Gusmão (2018, p. 3), citando Concklin (1987), destaca que problemas como a desorientação (*disorientation*) e o transbordamento cognitivo (*cognitive overhead*) são recorrentes nos diversos tipos de ambientes de informação digital. O autor também cita Oliveira e Vidotti (2013), que sugerem outros problemas de ordem navegacional, de

organização e de recuperação da informação que podem estar associados à questões qualitativas na Arquitetura da Informação desses ambientes. Partindo destas questões, e baseado naquilo que propõe a AI, começamos a pensar como seria organizada a estrutura do repositório para a promoção da recuperação da informação.

Decidir sobre como serão organizados os objetos digitais que compõem o repositório a ser criado para que não apenas eles estabeleçam uma relação entre si, mas também sejam localizados e recuperados, pode parecer uma tarefa fácil a primeira vista, mas não é. É preciso definir as categorias em que esses objetos serão classificados, como eles serão depositados e como serão estabelecidas estas relações, e a AI será essencial para auxiliar na tomada de decisão.

Como o repositório está sendo pensado

A etapa inicial foi nos reunirmos para responder a algumas questões básicas:

- Qual o objetivo do repositório? Reunir, armazenar, organizar, recuperar, preservar, disseminar.
- O quê? Produção científica e intelectual
- De quem? Docente, pesquisador, técnico, aluno
- Acesso? Livre
- Termos na caixa de busca: autor, ano, título, tema, assunto

Partindo destas questões, começamos a pensar como seriam organizados/estruturados os objetos digitais no repositório para que se realizasse a recuperação da informação. Pela familiaridade com o Dspace, ainda que este não tenha sido o escolhido para ser usado, a forma de pensar a organização foi hierárquica:

↳ Comunidades – contém as coleções e subcomunidades: DISCIPLINAS, CURSOS, FERRAMENTAS são algumas das comunidades sugeridas

↳ Coleções – o local onde se agrupam os documentos por tipo

Mas as comunidades e as coleções ainda não foram nomeadas.

A forma como se dará a organização dos objetos digitais para sua adequada recuperação foi pensada, inicialmente, conforme a imagem abaixo. Nela é possível compreender que dados e demais documentos (produção científica acadêmica, bibliográfica, vídeos, entrevistas em áudio, imagens, gráficos, tabelas), uma vez produzidos pelos atores sociais implicados serão todos encaminhados para o mesmo processo de curadoria de maneira a ser classificado como pertencente ao âmbito, ou não, das Humanidades Digitais. Em caso positivo este conjunto de dados, informações e documentos, passará pelo processo de depósito no repositório ao passo que serão duplicados em um servidor local como segurança.

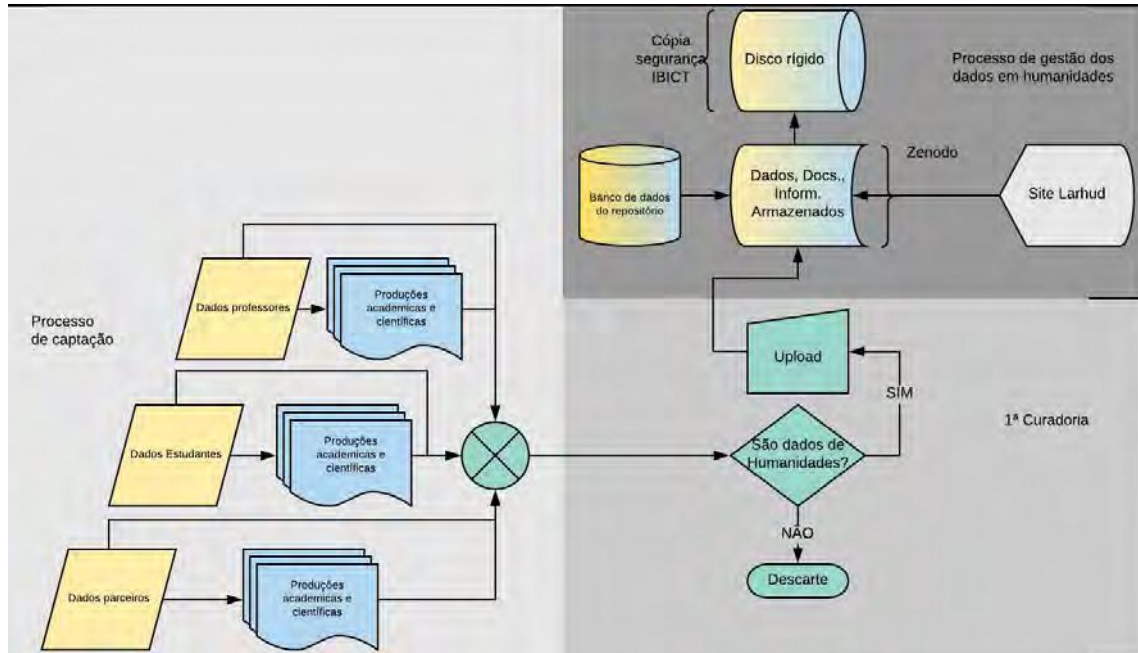


Figura 1: Diagrama de fluxo do repositório Larhud em construção.

Fonte: próprio autor, 2018.

A etapa seguinte foi identificar seus usuários: administradores, pesquisadores e possíveis visitantes; e para cada um deles foi definido um nível de permissões. Inicialmente, aos usuários comuns o acesso ao conteúdo do repositório se dará unicamente pelo sítio eletrônico do Larhud. Por esta entrada o usuário terá acesso ao conteúdo disponível e armazenado/preservado no Zenodo. Os membros do laboratório, no entanto, terão acesso tanto ao Zenodo como ao servidor local. A figura abaixo representa as permissões que serão dadas aos usuários:

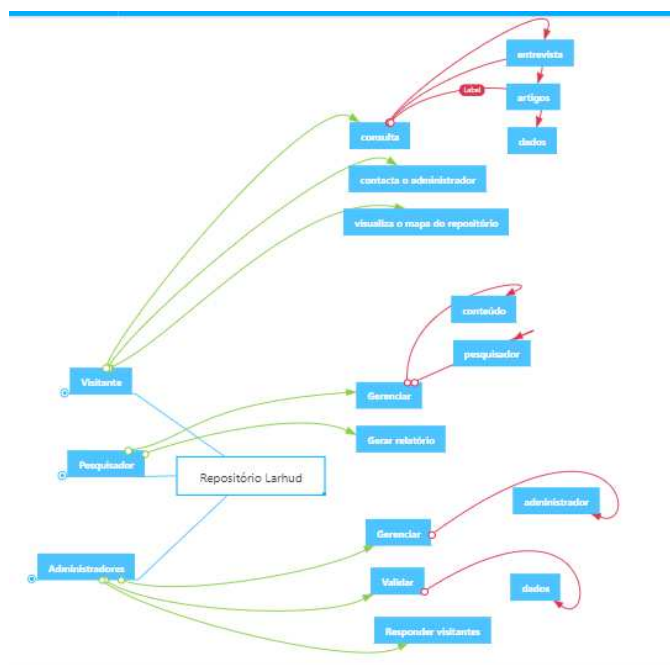


Figura 2: Permissões aos usuários do repositório
Fonte: próprio autor, 2018.

A etapa atual consiste na organização e representação da informação a partir da definição das comunidades e coleções que irão compor estruturalmente o repositório, bem como da seleção do tipo de metadados que serão utilizados.

Um problema que surgiu durante as discussões do grupo foi como pensar a interface para se acessar o acervo que estará disponível no Zenodo, ou seja, que compõe o repositório em si. Inicialmente essa etapa se dará acessando o site do Larhud, como dito acima. Procuramos diferentes pessoas que conhecem e trabalham com o Zenodo para saber se existe a possibilidade de combiná-lo com outro programa que permita a efetiva organização da informação como descrita até aqui, mas até o momento não conseguimos uma solução para este impasse. Essa questão é de suma importância pois precisamos garantir um acesso fácil, rápido e objetivo para um usuário que, enquanto potencial cientista das ciências humanas, pode não possuir competências digitais informacionais para uma procura profunda e complexa.

Considerações Finais

O momento atual, quando da escrita deste artigo, foi o de idealização, desenho da estrutura e busca por meios materiais de tornar concreto o projeto do repositório. O recurso do repositório em nuvem, apesar de seu servidor residir na União Européia, foi identificado como vantajoso para a implementação desse mesmo repositório destinado às humanidades.

A partir da proposta de criação de um repositório que atenda às Humanidades Digitais, o uso do Zenodo poderá permitir, no escopo das atividades realizadas pelo Larhud, armazenar os resultados de pesquisas científicas multitemáticas e disponibilizar o seu acesso a diferentes usuários, principalmente porque ele permite que a produção seja referenciada via um DOI próprio. Também esperamos que seja possível, a partir desse repositório, desenvolver meios para fortalecer uma comunidade em rede sobre a temática das Humanidades Digitais assim como espessar a produção científica sobre Humanidades Digitais à luz da Ciência da Informação.

Referências

CABEZAS, A. La Referencia and Cern enter into agreement to promote and expand open science in Latin America. LA Referencia [site]. 29 out. 2018. Disponível em: <<http://www.lareferencia.info/es/component/k2/item/218-la-referencia-and-cern-enter-into-agreement-to-promote-and-expand-open-science-in-latin-america>>. Acesso em: 29 set. 2018.

DACOS, M. **Manifesto das Digital Humanities**. That Camp Paris. [blog]. 2011. Disponível em: <<https://tcp.hypotheses.org/497>>. Acesso em: 29 set. 2018.

GUSMÃO, F. C. M. Elementos de arquitetura da informação no Repositório Eletrônico Institucional da UFPB. **Rev. Inf. na Soc. Contemp.**, Natal, RN, Número Especial, 2017. Disponível em: <<https://periodicos.ufrn.br/informacao/article/view/12282>>. Acesso em: 29 set. 2018.

INSTITUTO Brasileiro de Informação em Ciência e Tecnologia – IBICT. **Sobre Repositórios Digitais**. Disponível em: <<http://www.ibict.br/informacao-para-ciencia-tecnologia-e-inovacao%20/repositorios-digitais>>. Acesso em: 28 out. 2018.

MACEDO, F. L. O. **Arquitetura da Informação**: aspectos epistemológicos, científicos e práticos. Brasília: CID/UnB, 2005. 190 fl. (Dissertação de mestrado).



Bibliografia Consultada

YAGUI, M. M. M., VIEIRA, A. C. M., DA CRUZ, S. M. S. **Arquitetura e implementação de um repositório para plantas medicinais em uso no Brasil**. Disponível em: <<https://www.researchgate.net/publication/311841590> Arquitetura e implementacao d e um repositorio para plantas medicinais em uso no Brasil> . Acesso em: 13 ago. 2018.