

Declaração de Direito Autoral

Autores que submetem a esta conferência concordam com os seguintes termos:

a) Autores mantém os direitos autorais sobre o trabalho, permitindo à conferência colocá-lo sob uma licença [Licença Creative Commons Attribution](#), que permite livremente a outros acessar, usar e compartilhar o trabalho com o crédito de autoria e apresentação inicial nesta conferência.

b) Autores podem abrir mão dos termos da licença CC e definir contratos adicionais para a distribuição não-exclusiva e subsequente publicação deste trabalho (ex.: publicar uma versão atualizada em um periódico, disponibilizar em repositório institucional, ou publicá-lo em livro), com o crédito de autoria e apresentação inicial nesta conferência.

c) Além disso, autores são incentivados a publicar e compartilhar seus trabalhos online (ex.: em repositório institucional ou em sua página pessoal) a qualquer momento antes e depois da conferência. Fonte:

<https://enancib.ancib.org/index.php/enancib/xxivenancib/about/submissions#copyrightNotice>. Acesso em: 26 fev. 2025.

Referência

KAMA, Ana Flávia Lucas de Faria *et al.* Portal dos fóruns de Educação de Jovens e Adultos: migração, reorganização da informação e aplicação do software livre Tainacan. In: ENCONTRO NACIONAL DE PESQUISA E PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO, 24., 2024, Vitória. **Anais** [...]. [S. l.]: ANCIB, 2024. Disponível em: <https://enancib.ancib.org/index.php/enancib/paper/viewFile/2360/1505>. Acesso em: 26 fev. 2025.



24° ENANCIB
Encontro Nacional de Pesquisa em Ciência da Informação
Perspectivas Contemporâneas na Ciência da Informação
• Vitória - ES • Ancib • PPGCI/UFES



XXIV ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO – XXIV ENANCIB

ISSN 2177-3688

GT 8 – Informação e Tecnologia

PORTAL DOS FÓRUMS DE EDUCAÇÃO DE JOVENS E ADULTOS: MIGRAÇÃO, REORGANIZAÇÃO DA INFORMAÇÃO E APLICAÇÃO DO SOFTWARE LIVRE TAINACAN

PORTAL OF ADULT AND YOUTH EDUCATION FORUMS: MIGRATION, REORGANIZATION OF INFORMATION AND USING THE OPEN-SOURCE SOFTWARE TAINACAN

Ana Flávia Lucas de Faria Kama – Universidade de Brasília (UnB)

Michelli Pereira da Costa – Universidade de Brasília (UnB)

Dalton Lopes Martins – Universidade de Brasília (UnB)

Jéssica Bilac Gaspareto – Universidade de Brasília (UnB)

Modalidade: Trabalho Completo

Resumo: O Portal de Fóruns de Educação de Jovens e Adultos (EJA) é ambiente digital resultante de um conjunto de tecnologias sociais de articulação e difusão de informações que surgiu no início dos anos 2000. No entanto, diante dos avanços tecnológicos e das novas necessidades sociais e políticas do movimento EJA no país, tornou-se necessário reestruturá-lo. O objetivo deste trabalho é descrever o processo de migração do conteúdo do Portal para sua nova versão, construída a partir do software livre Tainacan. A metodologia utilizada compreende sete etapas, que vão desde a elaboração do diagnóstico do acervo existente até a abertura do acervo no novo sistema e sua disponibilização aos usuários. Os resultados alcançados até o momento incluem a migração dos dados para a nova versão do Portal, a criação e mapeamento de taxonomias e metadados específicos para a descrição dos objetos digitais, bem como o uso de extrações em massa de dados por meio de *scripts* em Python. Conclui-se que a aplicação do método de migração e disponibilização dos dados atingiu os resultados esperados, conferindo à nova versão do Portal a robustez necessária para abrigar o acervo de sua versão original.

Palavras-chave: Portal dos Fóruns de EJA; Tainacan; tecnologia social.

Abstract: The Adult and Youth Education (In Portuguese, EJA) Forums Portal is a social technology for the articulation and dissemination of information that emerged in the early 2000s. However, in light of technological advancements and the new social and political needs of the EJA movement in the country, it became necessary to restructure it. The objective of this paper is to describe the process of migrating the content of the Portal to its new version, built using the open-source software Tainacan. The methodology comprises seven stages, ranging from the preparation of a diagnosis of the existing collection to the opening of the collection in the new system and its availability to users. The results achieved so far include the migration of data to a newer version of the Portal, the creation and mapping of taxonomies and specific metadata for the description of digital objects, as well as the use of mass

data extractions through Python scripts. In conclusion, the application of the data migration and availability method has achieved the expected results, giving the new version of the Portal the necessary robustness to host the collection from its original version.

Keywords: Portal of Adult and Youth Education Forums; Tainacan; social technology

1 INTRODUÇÃO

A história da educação popular e pública no Brasil se confunde com a história dos movimentos sociais e com a busca da emancipação democrática e cidadã de parte da população brasileira discriminada de direitos (Fávero; Freitas, 2011; Freire, 1963). O trabalho e a luta pela Educação de Jovens e Adultos (EJA), envidados desde fins do século XIX e mais fortemente delineada na segunda metade do século XX, reforçam a necessidade de ações não só estatais, mas também sociais, na busca por uma perpetuação de política nacional e realização fecunda da EJA no país. Grandes esforços coletivos foram e são realizados diuturnamente por grupos de educadores e educandos de EJA de todo o Brasil. Um dos grandes produtos de suas ações foi a elaboração e manutenção do Portal dos Fóruns de EJA do Brasil, abrangendo os vinte e seis Estados e o Distrito Federal.

Pensado como um ambiente digital, os Portais de Fóruns de EJA datam do início dos anos 2000 e são espaços prioritariamente de articulação política dos atores regionais envolvidos nesta luta, muito embora reflitam também uma organização comunitária de informações, comunicações e conteúdos relacionados ao tema da EJA resultantes dos esforços intelectuais, sociais e políticos de todo o país na construção de uma educação popular gratuita, inclusiva e de qualidade (Machado, 2008; Portal dos Fóruns de EJA, 2024; Rodrigues; Bonfim; Ferreira, 2016).

Tecnologias e culturas digitais, como a aplicada no Portal, em especial aquelas que comunicam com o conceito de ciência aberta e cidadã, têm proporcionado ações relevantes nas últimas décadas no Brasil (Martins; Carvalho Júnior, 2017). Elas possibilitam a ideia da transmutação da memória digital de ações coletivas em uma prática social potente em transformações e mudanças políticas.

Ao partir dessas premissas, o objetivo deste trabalho é apresentar um estudo de caso das propostas e atividades realizadas no âmbito do Projeto Pegadas Paulo Freire (Rocha; Gomide, 2022), em especial em uma de suas metas, que diz respeito sobre permanência de um espaço vivo de produção e recriação de memória e história da educação de pessoas

trabalhadoras, ao descrever o processo de reestruturação, migração, organização e tratamento das informações contidas no Portal dos Fóruns de EJA para sua nova versão.

É importante destacar que o Portal, em sua versão original, foi concebido como um sistema de gerenciamento de conteúdo, tendo sido adotada a tecnologia Drupal¹ como *software* para sua implementação. Todo o processo de produção de conteúdo foi concebido originalmente de forma orgânica, utilizando-se de ferramentas de organização da informação tais como blog, páginas estáticas e diretórios de arquivos para ordenação da informação. Ao longo de aproximadamente 20 anos, esse processo orgânico resultou em diversos problemas relacionados à busca e recuperação da informação, dificuldade de navegação, além de estratégias de indexação de conteúdo divergentes e conflitantes. Esses desafios informacionais foram identificados ao longo deste estudo de caso, levando à decisão de adotar novas tecnologias capazes de migrar o conteúdo legado da plataforma original e implementar novos modelos de organização da informação. Assim, a escolha de uma tecnologia que integrasse um repositório digital e um sistema de gerenciamento de conteúdo para a web foi crucial para esta pesquisa.

A fim de realizar a reestruturação do Portal, vem sendo aplicada a metodologia de migração de dados à luz do modelo proposto por Martins, Carmo e Germani (2018), a qual agrupa sete etapas de análise de dados, passando por coleta, tratamento, migração, validação, dentre outras. Para tanto, foi utilizada a plataforma livre de repositório para acervos digitais, o Tainacan², em conjunto com o gestor de conteúdo para internet, o WordPress. Logo, este estudo apresenta uma análise do processo de reestruturação do Portal dos Fóruns de EJA do Brasil, ao detalhar os processos técnicos da migração de suas informações e o relacionamento entre si, assim a articulação de diferentes áreas do campo da Ciência de Dados e da Ciência da Informação durante esse percurso.

2 CONTEXTO SOCIAL E INSTITUCIONAL

O Portal dos Fóruns de EJA foi criado em 2005, e é formado por 27 sites representando cada estado (Fóruns Estaduais), o Distrito Federal (Fórum Distrital) e a esfera nacional (Fórum EJA Brasil). Esses sites são ambientes educativos, virtuais, interativos e multimídias,

¹ Sobre o *software* Drupal, ver informações em: <https://www.drupal.org/> (acesso em 26.06.2024).

² Mais informações sobre o Tainacan em: <https://tainacan.org/> (acesso em 26.06.2024).

desenvolvidos pela comunidade, acessíveis em: <http://forumeja.org.br/>. As páginas utilizam tecnologia LAMP baseada em CMS de código aberto, construídas na plataforma gratuita e código aberto Drupal, com atividades contínuas de pesquisa-ação para atualização e descoberta de novas ferramentas na comunidade de desenvolvedores de *software* livre. As páginas incorporam a metodologia de estudo do meio e utilizam a tecnologia com foco na educação de jovens, adultos e idosos trabalhadores. A construção e hospedagem do Portal são coordenadas pela Universidade de Brasília (UnB) e 27 administradores estaduais e distritais, destacando uma estrutura descentralizada, com apoio de vários segmentos que defendem os direitos da população EJA (Santos; Sousa, 2023).

O projeto Memória e História da Educação de Jovens, Adultos e Idosos Trabalhadores – Pegadas de Paulo Freire, foi proposto em 2022, por um conjunto de professoras e professores pesquisadores da UnB, como uma das ações de celebração do centenário do nascimento de Paulo Freire, tendo como objetivo,

[...] mapear, organizar e disponibilizar a produção no campo da memória e história vividas e vivas da educação voltada a pessoas trabalhadoras. Pretende se constituir como projeto em rede, agregando pesquisadores de instituições públicas de ensino, pesquisa e extensão, com vistas a sistematizar e produzir conhecimentos acerca da educação de pessoas trabalhadoras e da educação popular, para contribuir na luta em defesa da educação como direito humano. (Rocha; Gomide, 2022, p. 2).

Cabe ressaltar o entendimento do projeto como uma ação de preservação da memória do Portal e, mais do que isso, de apropriação de uma nova tecnologia que permita novos modos de organização da informação que podem se desdobrar em novas formas de uso da informação pela comunidade e pelos coletivos engajados na construção política da EJA. O uso estratégico da memória em novas perspectivas tecnológicas não é apenas uma forma de dar acesso ao passado, mas também de instalar o futuro em novos usos possíveis da informação. Como nos alerta Assmann (2011),

[...] O arquivo não é somente um repositório para documentos do passado, mas também um lugar onde o passado é construído e produzido. Essa construção não depende apenas de interesses sociais, políticos e culturais, mas é essencialmente codeterminada pelos meios de comunicação e pelas técnicas de registro (Assmann, 2011, p. 25–26).

Nesse sentido, o Portal realiza-se também como uma tecnologia social. Para Dagnino (2014), a tecnologia social promove a inclusão social e está ligada a produção de bens e serviços por meio de empreendimentos autogestionários. Ela é aplicada pelo produtor direto,

sem estruturas amplamente hierarquizadas e controles coercitivos. Aderente aos princípios da tecnologia social discutidas por Dagnino, o Portal desenvolve-se a partir de esforços colaborativos entre os grupos parceiros dos Fóruns de EJA e promove a autonomia e o domínio da tecnologia por parte sujeitos envolvidos com o projeto.

3 DIAGNÓSTICO DA VERSÃO ORIGINAL DO PORTAL DO FÓRUM DE EJA

Inicialmente, para o desenvolvimento da pesquisa, foi realizado um diagnóstico do Portal em sua primeira versão (v. 1), seguindo os componentes da arquitetura da informação, conforme a abordagem proposta por Rosenfeld, Morville e Arango (2015) no intuito de atender ao objetivo do projeto de pesquisa cuja finalidade é reestruturar o Portal dos Fóruns de EJA em torno de um novo modelo informacional que facilite a navegação, busca, recuperação e indexação do conteúdo produzido e arquivado pela comunidade.

Rosenfeld, Morville e Arango (2015) definem arquitetura da informação como a combinação dos esquemas de organização, rotulagem, busca e navegação em ambientes digitais como os *websites* e intranets. Os autores mencionam que a arquitetura da informação para *web* é composta por sistemas estruturados e interdependentes que são utilizados para organizar as informações disponíveis nas páginas *web* e de acordo com eles:

- o **sistema de organização** tem como finalidade distribuir o conteúdo no website utilizando as formas de organização através de esquemas e estruturas;
- o **sistema de navegação** permite uma interação do website com o usuário favorecendo caminhos que facilitam a obtenção da informação procurada;
- o **sistema de rotulagem** possui informações que podem ser acessadas através de palavras ou ícones que auxiliam na identificação do conteúdo;
- o **sistema de busca** é responsável pelo acesso rápido à informação contida no website.

O diagnóstico realizado é um trabalho de análise descritiva que identifica as configurações da arquitetura da informação das múltiplas páginas do Portal do Fórum EJA. Ele foi estruturado conforme a organização do Portal que apresenta segmentação nacional, estadual e distrital. Essa segmentação é representada por 26 páginas de fóruns estaduais, uma distrital e uma nacional, totalizando 28 ambientes digitais com informações relacionadas

aos debates da EJA de cada localidade. Apresenta-se a seguir de forma resumida alguns resultados que foram fundamentais para o estudo de caso em questão.

Após analisar os componentes de arquitetura da informação do Portal de Fóruns de EJA sob a perspectiva de Rosenfeld, Morville e Arango (2015) foi possível identificar as características principais e as especificidades das segmentações do Fórum EJA em relação aos critérios de organização, navegação, rotulagem e busca. Nota-se que, apesar de apresentar uma estrutura visual padronizada em relação ao estilo de organização e representação da informação, os diferentes segmentos apresentam nível de organização, representação e atividade distintos.

Em relação aos critérios de organização e navegação, observar-se que não há uso de vocabulário controlado. Notam-se diversas estratégias para rotular, organizar e representar a informação. Por um lado, há autonomia e liberdade por parte do administrador do Portal. Entretanto, há conflitos em relação ao acesso, pois a linguagem usada para rotular a informação não é padronizada e isso impacta na experiência de uso do usuário gerando um acesso por vezes confuso, as informações aparecem repetidas em níveis de categorias e formatos de organização e representação da informação, além de apresentar ambiguidades em relação aos termos e conceitos associados que dificultam o acesso à informação no Portal.

A minoria das páginas do Portal estava atualizada e mantinha uma rotina de publicação no momento do diagnóstico. Os portais: RS, SC, PA, AL, BA, PB, RR, PI, MT, AM, PE, MS, AP, SE, AC, PR, RN, RJ, CE e MA categorizaram a informação com poucos ou em alguns casos nenhum “menu”. Enquanto os portais: TO, GO, BR, MG, DF, ES, RO e SP categorizaram a informação com uso expressivo de abas e subseções para categorizar e segmentar a informação através de “menu”. Nota-se que não há padrão em relação a hierarquia das categorias pois elas variam de acordo com os elementos de organização e representação da informação nos Portais EJA. No entanto, ao adotar o padrão “menu” vinculado a seções a informação é representada por trilha que marca o caminho de acesso durante a navegação. De modo geral, verifica-se que independente do Portal adotar o uso do formato “menu” e/ou bloco lateral para organizar e representar a informação, o usuário encontra dificuldades de acesso à informação devido à falta de padronização e definição de vocabulário controlado que impacta na experiência de uso e acessibilidade do usuário, uma vez que, as informações são organizadas e representadas com camadas que dificultam o acesso.

4 METODOLOGIA

A metodologia utilizada para a reestruturação do Portal dos Fóruns de EJA possui semelhanças com a proposta por Martins, Carmo e Germani (2018), com exceção da fase de enriquecimento, que versa sobre uma aplicação semântica sobre os dados e que ainda não foi executada no âmbito do projeto. A proposta metodológica utilizada é composta de sete etapas que vão desde a elaboração do diagnóstico do acervo existente até o momento em que o acervo é aberto no novo sistema e disponibilizado aos usuários (Figura 1).

Figura 1 – Metodologia de migração e difusão dos acervos do Portal dos Fóruns de EJA



Fonte: Adaptado de Martins, Carmo e Germani (2018) e Oliveira e Martins (2019)

5 RESULTADOS PRELIMINARES

A reestruturação do Portal dos Fóruns de EJA, constante de uma das metas do Projeto Pegadas Paulo Freire, passou por uma primeira etapa de diagnóstico e reconhecimento, durante o ano de 2022 e 2023. Nesse momento, identificou-se a organização e arquitetura da informação contidas no site forumeja.org.br, desenvolvido com a ferramenta Drupal. Com o objetivo de migrar todos os itens digitais constantes da v.1 do Portal para uma nova instalação (v. 2) e domínio, elaborado em WordPress e com a ferramenta Tainacan, foram perpassadas todas as fases metodológicas compreendidas em: análise, coleta, tratamento, migração, validação e formação, difusão e colaboração em rede. Muito embora o projeto ainda não esteja finalizado e algumas dessas etapas continuem em desenvolvimento e reformulação, em

algum momento elas fizeram parte do método pelo qual vem se desenhando e executando a migração do Portal.

5.1 Análise: características técnicas do acervo

Para a realização da fase de análise, levantou-se informações junto aos gestores técnicos da v.1 do Portal, assim como um estudo *in loco*, a fim de averiguar como se encontrava a organização dos objetos digitais dentro do site. De forma resumida, os resultados dessa análise foram apresentados na seção 03 deste artigo. Como já dito anteriormente, é interessante lembrar que o Portal dos Fóruns de EJA é uma interface social e política utilizada como articulação de um movimento nacional em prol da educação popular de jovens e adultos, ou seja, é um esforço coletivo e muitas vezes voluntário de uma gama de pessoas que nem sempre são especialistas ou possuem experiência com organização, comunicação ou arquitetura da informação. Além disso, como já exposto, a ordenação do Portal é compartimentada de acordo com as estruturas federativas do Brasil, ou seja, havia 28 ambientes digitais com informações relacionadas aos debates da EJA.

Os ambientes digitais identificados eram compostos de páginas de conteúdo e objetos digitais referenciados. Para rastrear e reconhecer todas as páginas do Portal foi utilizada a ferramenta XML Maps Site³. A partir da estratégia foi possível identificar um montante de cerca de 442 páginas ou subpáginas, onde observou-se 2.654 objetos digitais pulverizados por várias subpáginas e depositados em forma de links carregados e, em sua grande maioria, sem informações identificadoras, como metadados estruturados ou descrições (Tabela 1).

Tabela 1 – Quantidade de objetos digitais extraídos do site forumeja.org.br

Região do Brasil	Quantidade
Brasil / Nacional	867
Norte	37
Nordeste	41
Sudeste	91
Sul	99
Centro-Oeste	819
Goiás AtoM*	700
TOTAL	2.654

Fonte: Elaborado pelos autores

*Todos os itens no formato .pdf

³ Para mais informações, ver em: <https://www.xml-sitemaps.com/> (acesso em 26.06.2024).

5.2 Coleta dos metadados e mídias

A partir da identificação de objetos digitais e suas respectivas URLs na v. 1 do Portal, iniciou-se a coleta de metadados e das mídias em três contextos distintos. Para isso, foi necessário criar um *script* em Python que permitisse a extração e raspagem em massa dos itens digitais, assim como a confecção de um segundo *script* para colher os vídeos dessa versão do Portal, mas hospedados na plataforma proprietária YouTube. O terceiro contexto de coleta derivou de uma demanda específica do projeto, envolvendo a extração de objetos descritos em uma instalação da plataforma AtoM⁴, do Centro de Memória Viva do Goiás (CMV-GO)⁵, relacionados a itens do Fórum de EJA Goiano (FEJA-GO). A Tabela 1 apresenta a quantidade de documentos extraídos a partir das estratégias mencionadas, discriminados por regiões do país.

Inicialmente, desenvolveu-se um *script* para raspagem de dados que, utilizando a linguagem Python, definia os metadados principais a serem extraídos por meio de um arquivo XML. A estrutura desse código⁶ empregava bibliotecas como Pandas, e ao fornecer a URL do XML e definir os metadados, o *script* convertia os dados em um arquivo CSV principal para *download*. No contexto da extração de vídeos⁷, também se utilizou Python, com as bibliotecas 'os' e 'urllib.request'. Uma lista com os *links* dos vídeos foi fornecida para que fossem baixados automaticamente. Esse processo garantiu a obtenção eficiente dos vídeos.

No terceiro contexto, utilizou-se parte da estrutura do primeiro *script*, mas com um diferencial significativo: a utilização do AtoM e de metadados do Dublin Core para gerenciar e baixar os dados. Por meio da listagem de verbos do AtoM, foram catalogados todos os metadados provenientes dos objetos do CMV-GO e do FEJA-GO. Um detalhe importante é que o *script* incluía uma funcionalidade para conciliar automaticamente o nome do arquivo com o *download* de imagens ou PDFs relacionados a cada item, criando uma coluna no CSV. O

⁴ O AtoM é um software livre e com o código-fonte aberto, desenvolvido para ser uma interface de acesso aos documentos arquivísticos. Fonte: <https://www.atom.unb.br/about>, acessado em 13.09.2024.

⁵ O Centro de Memória Viva (CMV-GO) é um Centro de Referência de EJA, educação popular e movimentos sociais, por meio de pesquisa documental, história oral e registro de experiências articulado a outros estados. Fonte: <https://ppge.fe.ufg.br/> (acesso em 13.09.2024).

⁶ Disponível em: <https://github.com/jbilacg/OAI-PMH-Downloader> (acesso em: 27.06.2024).

⁷ Disponível em: <https://github.com/jbilacg/Video-Downloader-from-YouTube> (acesso em: 27.06.2024).

objetivo era que, ao serem carregados no Tainacan, esses itens já estivessem associados aos seus metadados descritivos correspondentes.

Assim, cada um dos contextos de coleta de dados foi meticulosamente planejado e executado para garantir a integridade e a organização dos dados extraídos, facilitando sua posterior utilização e integração na plataforma Tainacan (Tabela 2).

Tabela 2 – Origem e tratamento dos dados extraídos e importados

Origem dos dados	Processo de exportação	Situação
Objetos digitais da v.1 do Portal: páginas e subpáginas	Script de linguagem Python	100% tratados 96,2% importados
Links de vídeos do YouTube mapeados na v.1 do Portal	Script de linguagem Python	100% tratados 100% importados
Objetos digitais e suas descrições no AtoM do Centro de Memória Viva (FEJAGO*)	Script de linguagem Python	100% tratados e importados, mas, passam por processo de revisão

Fonte: Elaborado pelos autores
*Fórum de EJA Goiano

5.3 Tratamento e refinamento dos metadados

A partir da coleta dos dados e metadados identificou-se que as mídias referenciadas nas páginas do Portal da v. 1 não estavam descritas satisfatoriamente. A coleta automática dos metadados permitiu identificar apenas as informações referentes aos títulos dos arquivos, formatos e tamanhos das mídias.

A etapa de tratamento teve sua primeira fase realizada durante o período de 2022 e 2023, contudo, segue com atividades em constantes atualizações. Nesta primeira fase, iniciou-se o tratamento dos dados já nas planilhas e *scripts* de exportação dos três contextos distintos de coleta explanados anteriormente, com refinamento e padronização de metadados. O processamento dos dados levou em consideração o mapeamento e a estrutura de metadados proposta pela equipe do projeto, que passou por diversas alterações e adaptações de forma a refletir o conjunto de necessidades de informação do movimento da EJA, assim como de representar o mais fiel possível os objetos digitais coletados. Atualmente, um conjunto de 16 metadados estão organizados, conforme indicado no Quadro 1.

**XXIV Encontro Nacional de Pesquisa em Ciência da Informação – XXIV ENANCIB
Vitória-ES – 04 a 08 de novembro de 2024**

Quadro 1 – Metadados da versão 2 do Portal dos Fóruns de EJA

Nome do metadado	Tipo de metadado	Mapeamento dos metadados - Dublin Core
Título	Texto simples	DC Elements: dc.title
Autoria	Taxonomia (livre)	DC Elements: dc.creator
Assunto	Taxonomia	DC Elements: subject
Outros assuntos	Taxonomia (livre)	-
Ano de publicação	Taxonomia	DC Elements: dc.date
Data de publicação	Texto simples	DC Terms: dc.created
Tipo de documento	Taxonomia	DC Elements: dc.type
Coleção ou evento	Taxonomia	DC Elements: dc.relation
Outra coleção ou evento	Texto simples	-
Local de publicação	Taxonomia	DC Elements: dc.coverage
Observação sobre o local	Taxonomia (livre)	DC Terms: dc.spatial
Editor(a) ou Instituição	Taxonomia (livre)	DC Elements: dc.publisher
Descrição	Descrição principal	DC Elements: dc.description
Notas	Texto longo	-
URL	URL	DC Elements: dc.identifier
Licença	Taxonomia (campo oculto)	DC Elements: dc.right

Fonte: Elaboração dos autores – junho/2024

O esquema de metadados adotado no Portal v.2 foi composto por metadados classificados em quatro tipos, a saber: taxonomia, texto simples, texto longo e URL. A adaptação desse esquema deu-se em função da diversidade tipológica dos documentos armazenados e disponibilizados no Portal. Ressalta-se que parte considerável do acervo é composto por material não bibliográfico, dos quais evidencia-se a dificuldade em localizar informações descritivas, tais como relacionadas a imprensa.

Considerando os limites acima descritos, o esquema de metadados buscou parâmetros flexíveis, de modo a se adequar aos diferentes tipos de documentos e que pudesse ser compreensível para os diferentes tipos de usuários catalogadores no sistema. Dessa forma, grande parte dos metadados são de uso não obrigatório e baseados em taxonomias, conforme descrito na Tabela 3.

Tabela 3 – Taxonomias criadas x situação

Taxonomia	Descrição	Quantidade de termos	Situação
Ano de Publicação	Ano(s) e/ou década(s) de publicação ou elaboração do documento.	14 termos raiz 140 termos filhos	Tratada e finalizada
Assunto	Tesouro evoluído a partir dos termos incluídos nas primeiras catalogações e com base no Tesouro brasileiro de educação.	50 termos raiz 114 termos filhos	Primeiro tratamento finalizado - em construção
Outros assuntos	Assuntos que não estão dispostos na taxonomia “Assuntos”, criada para dar subsídios ao Tesouro de assuntos	2029 termos raiz	Primeiro tratamento culminou na versão preliminar do Tesouro de Assuntos
Autoria	Taxonomia livre das autorias dos documentos	937 termos raiz	Não tratada
Coleção ou evento	Indicação de coleções e/ou eventos nos quais o documento diz respeito	118 termos raiz	Primeiro tratamento finalizado - em construção
Editor(a) ou Instituição	Editor(a) ou instituição responsável pela publicação ou veiculação do documento	234 termos raiz	Não tratada
Licença	Indicação da licença de distribuição gratuita na internet (<i>creative commons</i>)	0	Não tratada e oculta
Local de publicação	Local onde o documento foi publicado ou divulgado	10 termos raiz 5602 termos filhos	Tratada e finalizada
Observações sobre o local	Informações sobre o local, como bairro, região, dentre outros	18 termos raiz 1 termo filho	Primeiro tratamento finalizado - em construção
Tipo de documento	Tipologias documentais dos objetos digitais	152 termos raiz	Não tratada
TOTAL	-	3562 termos raiz 5857 termos filhos	-

Fonte: Elaborado pelos autores - dados extraídos em 17.06.2024

5.4 Migração dos metadados e dos objetos digitais

A migração desses itens passou por duas fases distintas. A primeira diz respeito aos dois primeiros contextos de coleta de itens digitais da v.1 do Portal: extração de itens diretamente das páginas mapeadas da v.1 e àquela realizada diretamente da plataforma YouTube, ambas por *scripts* em linguagem Python, as quais resultaram em uma identificação bastante básica de metadados desses objetos, no geral, apenas o título, extensão e tamanho do arquivo, além de sua URL de origem. Ou seja, um desafio foi colocado diante da migração e importação desses itens para a v.2 Portal: não seria possível realizar a importação sem auxílio humano para descrição dos metadados. Portanto, para que essa nova versão chegasse em uma fase de uso e testes robustos, foi necessário elaborar uma metodologia de submissão desses itens no repositório da v.2 do Portal. Os 1.954 objetos extraídos, tanto diretamente da v.1 quanto do YouTube (Tabela 1), foram organizados em uma planilha de controle na qual vários colaboradores do projeto, estagiários supervisionados de graduação em

Biblioteconomia da UnB e integrantes dos Fóruns de EJA têm participado ativamente da submissão desses itens, após receberem breves treinamentos e estudarem os tutoriais de submissão produzidos por demais membros da equipe do projeto (Rocha, 2024)⁸.

Já a segunda fase da migração dos itens passou pela coleta e tratamento daqueles documentos identificados e extraídos do AtoM do CMV-GO. Tais itens já possuíam uma prévia descrição e catalogação, com importantes metadados estabelecidos. A partir do XML mapeado em Dublin Core da coleção do AtoM, referentes a documentos do FEJA-GO, foram realizados tratamentos dentro do OpenRefine de forma a organizar os dados extraídos em colunas correspondentes aos metadados (Quadro 1) do repositório da v.2 do Portal. Após esse tratamento, os dados, junto com seus respectivos objetos digitais, foram transportados para a coleção do Goiás, a partir das ferramentas de importação do Tainacan. A reunião das três estratégias resultou na quantidade de 2.443 itens migrados, o que representa 96,2% do total.

5.5 Validação e formatação dos metadados

O procedimento de validação e formatação vem sendo executado desde que a primeira migração foi realizada dentro do repositório da v.2 do Portal. Muitas alterações já foram propostas, tanto pela equipe do projeto, quanto pelos integrantes dos segmentos dos Fóruns de EJA. As principais mudanças refletiram-se na exclusão, inclusão e/ou alteração de metadados; no tratamento e criação de vocabulários controlados para algumas taxonomias; criação e alteração de filtros de busca; mapeamento de metadados e melhorias no SEO da instalação, de forma a impulsionar a descoberta das informações do repositório; no uso de plugins do WordPress que facilitassem a segurança e gestão das informações; as consequências do uso do repositório na rotulagem dos dados e no design no Portal em sua v.2; em formas de submissão dos itens de coleta sem metadados previamente definidos; assim como nas constantes melhorias de seu funcionamento.

Parte da validação dos dados também foi realizada na apresentação da v.2 do Portal para coordenações de Fóruns de EJA de cada Estado do país e do DF durante a 1ª Oficina Nacional de Reestruturação do Portal dos Fóruns de EJA do Brasil, que aconteceu nos dias 26 e 27 de abril de 2024, em Brasília.

⁸ Durante 2022 e 2023, elaborou-se manual de submissão e uso do repositório do Portal em sua versão 2, em Tainacan.

5.6 Difusão do acervo e promoção da colaboração em rede

A etapa de difusão e colaboração em rede é realizada em consonância e concomitante com o processo de validação e formatação da v.2 do Portal. Contudo, há a previsão para um lançamento oficial da v.2 ainda em 2024, durante o XVIII Encontro Nacional De Educação De Jovens e Adultos (ENEJA), que poderá resultar em uma difusão mais robusta, tanto do Portal em si quanto do repositório, com expectativas de culminar em novas colaborações em rede para melhorias constantes do projeto.

6 CONSIDERAÇÕES FINAIS

Os movimentos da EJA no Brasil buscam influenciar a construção de um projeto político de sociedade justo e democrático, passando pela reorganização de diversas estruturas sociais. A presença em ambiente digital é uma de suas vertentes de atuação e articulação, promovendo a história viva e instrumental para suas articulações. Nesse sentido, a atualização e otimização dos recursos do Portal, bem como a organização de seu conteúdo, é de fundamental importância para seus fins.

A reestruturação do Portal dos Fóruns de EJA do Brasil buscou adequar-se aos princípios organizativos e políticos dos grupos. Para tanto, a equipe do projeto contou com a colaboração de pesquisadores de diferentes áreas de conhecimento e níveis de formação. Em nível instrumental, foram utilizadas tecnologias livres, de maneira intencional e crítica. Em uma perspectiva ampla, o projeto de reestruturação buscou desenvolver uma tecnologia social, conforme discutido por Dagnino (2014). O método de reestruturação do Portal foi orientado pela metodologia proposta por Martins, Carmo e Germani (2018) e ampliada por Oliveira e Martins (2019). A metodologia mostrou-se pertinente e flexível aos objetivos aqui estabelecidos. Dos sete passos por ela indicados, foi possível implementar e discutir seis. Em especial, destaca-se a fase de organização e tratamento dos dados como elemento estruturante da proposta apresentada. Em síntese, ressalta-se as etapas de análise e coleta como fundamentais para a organização dos esforços empenhados. As etapas foram auxiliadas por instrumentos de rastreamento de URLs e *scripts* em Python. Já as etapas de tratamento e migração dos dados exigiram grandes esforços de representações descritivas e temáticas dos

objetos digitais selecionados, especialmente quando consideradas as necessidades de normalização de taxonomias. As últimas etapas da metodologia ainda estão em curso e em avaliação. A única etapa da metodologia desconsiderada pela experiência relatada versa o refinamento dos metadados com o auxílio de ontologias para um aprimoramento semântico dos dados. Sugere-se que esse limite da pesquisa seja extrapolado em estudos futuros.

As contribuições do estudo se destacam pela aplicação da metodologia de migração de dados, já discutida anteriormente, com o uso do *software* livre Tainacan, evidenciando como tecnologias abertas podem fortalecer iniciativas sociais populares. Além de estabelecer uma base teórica para futuros estudos sobre preservação e democratização da informação digital, a pesquisa apoia a divulgação e o fortalecimento do Portal dos Fóruns de Educação de Jovens, Adultos e Idosos trabalhadores do Brasil e dos Estados, ampliando sua visibilidade e promovendo o engajamento da comunidade em ações educacionais e sociais. Entre as limitações do trabalho, destacam-se a interoperabilidade entre ferramentas, especialmente na integração do Tainacan com outras plataformas, e a diversidade dos dados disponíveis, que apresentam pouca padronização, dificultando sua organização e migração eficiente. A implementação da nova versão do Portal junto à comunidade dos Fóruns de EJA e ao público em geral ainda está em fase inicial e será abordada em uma próxima etapa, com estratégias de treinamento e engajamento para garantir sua adoção.

REFERÊNCIAS

ASSMANN, A. **Espaços da recordação**: formas e transformações da memória cultural. Campinas, SP: Editora da Unicamp, 2011.

DAGNINO, R. **Tecnologia social**: contribuições conceituais e metodológicas. Florianópolis: Ed. Insular, 2014.

FÁVERO, O.; FREITAS, M. A educação de jovens e adultos: um olhar sobre o passado e o presente. **Revista Inter Ação**, v. 36, n. 2, 28 dez. 2011.

FREIRE, P. Conscientização e alfabetização — Uma nova visão do Processo. **Estudos Universitários**, v. 4, p. 5–23, 1963.

MACHADO, M. M. **Formação de educadores de jovens e adultos**: II Seminário Nacional. Brasília: Secad/MEC; UNESCO, 2008.

**XXIV Encontro Nacional de Pesquisa em Ciência da Informação – XXIV ENANCIB
Vitória-ES – 04 a 08 de novembro de 2024**

MARTINS, D.; CARVALHO JÚNIOR, J. M. C. Memória como prática na cultura digital. *In*: NÚCLEO DE INFORMAÇÃO E COORDENAÇÃO DO PONTO BR. **Pesquisa sobre o uso das tecnologias de informação e comunicação nos equipamentos culturais brasileiros: TIC Cultura 2016**. São Paulo: Comitê Gestor da Internet no Brasil, 2017. p. 45–52.

MARTINS, D. L.; CARMO, D. DO; GERMANI, L. G. 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.

OLIVEIRA, L. F. R.; MARTINS, D. L. Fundação Nacional de Artes: estudo de caso da migração e publicação dos dados do acervo digital com o software livre Tainacan. ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 20., 2019, Florianópolis. **Anais...** Florianópolis: ANCIB, 2019. Disponível em: <https://conferencias.ufsc.br/index.php/enancib/2019/paper/view/1352>. Acesso em: 17 jun. 2024.

PORTAL DOS FÓRUNS DE EJA. **Histórico dos Fóruns de EJA**. Disponível em: <http://forumeja.org.br/node/1191>. Acesso em: 17 jun. 2024.

ROCHA, E. N.; GOMIDE, C. S. **Projeto memória e história da educação de jovens, adultos e idosos trabalhadores - Pegadas Paulo Freire**. Brasília: [s.n.], 2022.

ROCHA, G. M. **Manual para descrição de itens no Tainacan**: Portal de Fóruns de EJA. Brasília: [s.n.], 2024. Disponível em: <https://desenvolvimento.forumeja.org.br/brasil/manual-para-descricao-de-itens-no-tainacan/>. Acesso em: 18 jun. 2024.

RODRIGUES, M. E. C.; BONFIM, R. G. M.; FERREIRA, K. H. H. F. Portal do Fórum Goiano de EJA como espaço formador na educação de jovens e adultos. **Anais do XI Encontro Estadual de História da ANPUHGO**, v. 1, n. 3, p. 879–886, 2016.

ROSENFELD, L.; MORVILLE, P.; ARANGO, J. **Information architecture: for the web and beyond**. 4. ed. Sebastopol, CA: O'Reilly Media, 2015.

SANTOS, L. S.; SOUSA, V. M. **Relatório do Projeto de Memória e História da Educação de Jovens, Adultos e Idosos Trabalhadores - Pegadas Paulo Freire: meta 2, reestruturação do Portal dos Fóruns de EJA - Diagnóstico do Portal v.1**. Brasília: [s.n.], 2023. Acesso em: 17 jun. 2024.

NOTA DE AGRADECIMENTOS E RECONHECIMENTO

Este estudo de caso não teria sido possível sem a contribuição de Liliane Simões dos Santos, Victória Marques de Sousa, Maria Cecília Costa de Sousa, Ana Beatriz Berto de Oliveira, Gabriela Melo Rocha, Rodrigo Freire de Oliveira, Suelane Silva Ramos dos Santos.