

PRÁTICAS DE DIGITALIZAÇÃO: experiências na construção de fontes para a História da educação matemática

DIGITIZATION PRACTICES: experiences in building sources for the History of mathematics education

Guilherme Rodrigues Magalhães¹

 ORCID iD: <https://orcid.org/0009-0002-9905-0915>

Danielle Assunção Silva²

 ORCID iD: <https://orcid.org/0000-0003-0125-745X>

Maria Célia Leme da Silva³

 ORCID iD: <https://orcid.org/0000-0001-6029-0490>

Submetido: 04 de julho de 2025

Aprovado: 29 de outubro de 2025

RESUMO

Este artigo apresenta a experiência do grupo de pesquisa do qual participam os autores na digitalização de livros didáticos de matemática, tratados como fontes para a História da educação matemática. Partindo do entendimento de que o fazer historiográfico tem início na produção de documentos a partir de determinados objetos, o texto compartilha um método de digitalização portátil e de baixo custo, utilizando um celular e um suporte com iluminação. Após a digitalização, são detalhadas as etapas de edição do arquivo, incluindo conferência de páginas, melhoria da qualidade das imagens e conversão em texto pesquisável. Antes do detalhamento prático, discute-se o uso de recursos digitais no fazer historiográfico, abordando a perda da materialidade dos livros didáticos e a necessidade de assegurar a autenticidade histórica das fontes digitais por meio de rigor metodológico. Também se descreve a estrutura de dois acervos de livros físicos visitados pelo grupo, nos quais se concentrou parte significativa do processo. Conclui-se que procedimentos simples e acessíveis podem ampliar o acesso de pesquisadores a fontes raras, sem circulação ou distantes, além de contribuir para a preservação dos documentos originais.

Palavras-chave: Livros didáticos; Metodologia de pesquisa; Preservação digital.

ABSTRACT/ RESUMEN/ RÉSUMÉ

This article presents the experience of the research group to which the authors belong in the digitization of mathematics textbooks, treated as sources for the History of mathematics education. Starting from the understanding that historiographical practice begins with the production of documents from specific objects, the text shares a portable and low-cost digitization method, using a mobile phone and a stand with lighting. Following digitization, the file processing steps are detailed, including page verification, image quality improvement, and conversion to searchable text. Before the practical details, we discuss the use of digital resources in historiographical practice, addressing the loss of materiality of textbooks and the need to ensure the historical authenticity of digital sources through methodological rigor. We also describe the structure of two physical textbook collections visited by the group, where a significant part of the process was concentrated. We conclude that simple and accessible procedures can broaden researchers' access to rare, out-of-circulation, or distant sources, in addition to contributing to the preservation of original documents.

Keywords: Textbooks; Research methodology; Digital preservation.

¹ Mestrando em Ensino de Matemática pelo Instituto de Matemática e Estatística da Universidade de São Paulo (IME-USP). Avenida Caxingui 231, ap. 41-B, São Paulo, SP, Brasil, CEP: 05579-000. E-mail: guilherme@ime.usp.br.

² Mestranda em Educação Matemática pelo Programa de Pós-Graduação em Educação Matemática da Universidade Estadual Paulista "Júlio de Mesquita Filho" (PPGEM-UNESP). Rua Francisco Marson 234, ap. 36-B, São Paulo, SP, Brasil. E-mail: danielle.assuncao@unesp.br.

³ Doutora em Educação (Currículo) pela Pontifícia Universidade Católica de São Paulo (PUC/SP). Professora Associada da UNIFESP, Diadema, SP, Brasil e Professora do Programa de Pós-Graduação em Educação Matemática da UNESP, de Rio Claro. Rua Rio Grande 551, ap. 11B, Vila Mariana, São Paulo, SP, Brasil, CEP: 04018-001. E-mail: celia.leme@unifesp.br.

1 Introdução

Em história, tudo começa com o gesto de separar, de reunir, de transformar em “documentos” certos objetos distribuídos de outra maneira. Esta nova distribuição cultural é o primeiro trabalho. Na realidade, ela consiste em produzir tais documentos, pelo simples fato de recopiar, transcrever ou *fotografar* estes objetos mudando ao mesmo tempo o seu lugar e o seu estatuto. Este gesto consiste em “isolar” um corpo, como se faz em física, e em “desfigurar” as coisas para constituí-las como peças que preencham lacunas de um conjunto, proposto a priori. [...] As origens de nossos Arquivos modernos já implicam, com efeito, na combinação de um grupo (os “eruditos”), de lugares (as “*bibliotecas*”) e de práticas (de *cópia*, de impressão, de comunicação, de classificação, etc.) (De Certeau, 2002, p. 81, grifos nossos).

O presente artigo tem como objetivo refletir e compartilhar práticas de um coletivo de pesquisadores, especificamente sobre processos de cópia com tecnologias digitais – que designamos por digitalização –, tratamento e organização de livros didáticos. Esse coletivo é o Grupo de Estudos e Pesquisas em Geometria Escolar: história e formação de professores (GEPGE), criado e cadastrado no Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) em 2023, sob coordenação da professora Célia Leme (terceira autora deste texto). O GEPGE congrega pesquisadores desenvolvendo trabalhos em níveis de mestrado, doutorado e pós-doutorado⁴.

Apesar de sua recente constituição, o grupo se dedica a investigações que têm como foco a produção de história da geometria escolar. Desde 2023, todos os projetos de pesquisa desenvolvidos no GEPGE adotam, em comum, os livros didáticos de matemática, pertencentes a diferentes épocas e segmentos de ensino, como objeto de investigação. Há estudos sobre os segmentos de ensino hoje nomeados por Anos Iniciais do Ensino Fundamental, Anos Finais do Ensino Fundamental e do Ensino Médio.

A líder do grupo acumulou uma longa trajetória de aprendizagem e experiência por meio de sua participação no Grupo de Pesquisa de História da Educação Matemática (GHEMAT), de 2005 a 2022, tendo contribuído com a inserção de fontes no Repositório Digital Institucional da Universidade Federal de Santa Catarina (UFSC), vinculado ao GHEMAT-Brasil.

Com a criação do GEPGE e o desenvolvimento de novos projetos de pesquisa, e com a chegada de estudantes oriundos de diferentes instituições e regiões⁵, tornou-se necessário fomentar um espírito coletivo, construindo novas aprendizagens diante dos desafios impostos

⁴ O GEPGE é atualmente constituído por: Maria Célia Leme da Silva (UNIFESP/UNESP), Ana Paula Jahn (IME-USP), Guilherme Rodrigues Magalhães (Mestrando/USP), Ederson Sales Pastor (Doutorando/UNESP/Bauru), Josilaine Aparecida Pianoschi Malmonge (Doutoranda/UNESP/Bauru), Talita Moreira (Doutoranda/UNESP/Rio Claro), Danielle Assunção da Silva (Mestranda/ UNESP/Rio Claro) e Guilherme de Souza Silva (Mestrando/UNESP/Rio Claro).

⁵ O GEPGE contempla pesquisas desenvolvidas no Mestrado Profissional em Ensino de Matemática do IME-USP, no Programa de Pós-graduação em Educação para Ciência da Unesp de Bauru e no Programa de Pós-graduação em Educação Matemática da Unesp de Rio Claro.

pela produção científica. Para esses estudantes, o acesso às fontes de pesquisa, especialmente para aqueles residentes no interior dos estados, nem sempre é simples, visto que acervos e bibliotecas, em geral, localizam-se em grandes centros.

Nas últimas décadas, as tecnologias digitais passaram a ocupar um espaço significativo nas práticas de pesquisa. Diversos estudos na área de História têm mobilizado recursos digitais, porém, frequentemente sem discutir as especificidades teórico-metodológicas de sua utilização, seus impactos e, em muitos casos, sem sequer serem referenciados nas principais publicações da área (Brasil; Nascimento, 2020).

O GEPGE também pode ser incluído como um espaço, que apenas recentemente, motivado pelas experiências que relatamos neste artigo, passou a considerar a temática da digitalização como objeto de discussão. Nesse movimento, tem-se evidenciado a necessidade e a importância de detalhar os procedimentos de digitalização como parte relevante dos aspectos teóricos-metodológicos das pesquisas.

Na História da educação matemática⁶, observa-se igualmente uma crescente utilização de acervos e arquivos digitais, assim como de fontes digitalizadas. Um exemplo é o Repositório de Conteúdo Digital da UFSC, organizado pelo GHEMAT-Brasil e disponível desde 2012:

Esses novos espaços de guarda não só ampliam as possibilidades de investigação e potencializam novas problemáticas, como também exigem uma adaptação das metodologias e de práticas de análise crítica das fontes disponíveis, bem como atenção às limitações de uso a partir de direitos autorais, organização, seriação, capacidade, compatibilidade e armazenamento (Gregório; Costa; Maciel, 2025, p. 2).

A digitalização de materiais tem ampliado significativamente o acesso às fontes de pesquisa historiográfica por parte dos pesquisadores que se encontram distantes de acervos e bibliotecas físicas⁷. Além disso, essa prática possibilita novas formas de interação com documentos históricos.

Assim como Santos (2021), entendemos que as fontes digitais podem ser divididas entre as nativas digitais – aquelas produzidas diretamente em formato digital – e digitalizadas – aquelas convertidas de suportes físicos. Ambas oferecem vantagens importantes, como o acesso remoto, a preservação dos documentos originais e a possibilidade de análises em larga escala. No entanto, também impõem desafios metodológicos relevantes, como a perda da materialidade dos documentos e os riscos de interpretação mediados pelas tecnologias digitais.

⁶ Usamos a mesma distinção feita por Valente (2013) e grafamos *Educação Matemática* como referência ao campo acadêmico, responsável pelas investigações sobre ensino e aprendizagem de matemática, e grafamos *educação matemática* para fazer referência aos processos de ensino e aprendizagem da Matemática desde tempos imemoriais.

⁷ No GEPGE, temos estudantes que residem em Birigui, Tibiriçá, Jundiá, Rio Claro, Pindamonhangaba, cidades do interior do estado de São Paulo e em São Paulo, capital.

Em consonância com o artigo de Santos (2021), reconhecemos a relevância de Arquivos como a Hemeroteca Digital Brasileira, a Biblioteca Nacional Digital e o Centro de Pesquisa e Documentação de História Contemporânea na consolidação das práticas de pesquisa em ambiente digital. Esses acervos oferecem ferramentas de busca por palavras-chave e disponibilizam documentos digitalizados na íntegra, possibilitando pesquisas que, antes, exigiam deslocamentos extensos e presença física nas instituições.

Embora não seja objetivo deste texto discutir ou analisar a produção de fontes digitais de pesquisa, destacamos a pertinência e a necessidade de refletir sobre a digitalização de fontes e seus impactos no trabalho do historiador. Brasil e Nascimento (2020, p. 201) chamam a atenção para os cuidados inerentes ao processo de digitalização de documentos:

Quando um registro histórico – seja ele um manuscrito, uma carta, uma edição de jornal, uma foto, um livro etc. – converte-se, por meio de algum processo computacional, em um documento digital, ocorre aí uma mudança que dificilmente poderia ser considerada trivial. Apesar de a informação contida na fonte continuar “sendo a mesma” – no sentido de que a digitalização não alteraria substancialmente o conteúdo do registro histórico –, podemos dizer que a modificação na “materialidade” da fonte histórica nos conduz, inevitavelmente, a uma nova condição em relação ao modo de lidarmos com a informação ali contida.

Essa nova materialidade precisa ser compreendida como uma cópia digitalizada – isto é, uma reprodução de um objeto físico transformada por tecnologias – e, como tal, para que seja considerada *verdadeira*, no sentido de fidedigna ao original, exige um rigoroso controle de qualidade. É necessário que a procedência e o contexto da digitalização assegurem a integridade do processo, garantindo que todo o conteúdo do material esteja presente na versão digital.

Uma característica relevante desse novo suporte é o desaparecimento, parcial ou total, de propriedades, como cor, brilho, luz, textura, maciez, que podem ser determinantes para a interpretação de determinadas fontes históricas (Brasil; Nascimento, 2020). Desta maneira, consideramos fundamental deixar explícito, nos procedimentos metodológicos da pesquisa, o uso de fontes digitais, detalhando como se deu o processo de digitalização, quem o realizou, bem como a forma como os arquivos foram organizados e disponibilizados.

Outro aspecto a ser considerado é que a cópia digitalizada adquire o caráter de reprodutividade, podendo ser reproduzida indefinidamente. A possibilidade de acesso ao registro histórico amplifica-se, ou, para usarmos um termo muito em voga, ele *viraliza*. Isso nos conduz ao outro aspecto da mudança de materialidade: ao ser digitalizada, a fonte digital

permite a inclusão de metadados⁸ que, juntamente com o próprio texto digitalizado, pode ser pesquisável.

Em síntese, em conformidade com os cuidados, alertas e destaques apresentados anteriormente, reafirmamos a necessidade de um trabalho atencioso por parte dos pesquisadores no processo de produção de fontes digitais, com o mesmo rigor metodológico exigido no tratamento de fontes não digitais. No caso do nosso grupo, a demanda consistiu em transformar livros didáticos antigos em fontes digitais, a fim de viabilizar o acesso a esses materiais por pesquisadores que não tinham condições de consultar presencialmente as bibliotecas ou os acervos onde os exemplares estavam armazenados.

Os livros didáticos digitalizados pelo grupo referem-se a diferentes níveis de ensino, em conformidade com os projetos em desenvolvimento, ou seja, livros destinados ao antigo grupo escolar (atual Anos Iniciais do Ensino Fundamental), ao curso ginásial (atual Anos Finais do Ensino Fundamental) e ao curso colegial (atual Ensino Médio). O período de publicação dos livros abrange, até o presente momento, os anos 1930 até 1970, porém temos a intenção de chegar até o final do século XX. A escolha dos livros didáticos digitalizados foi feita por cada pesquisador, em consonância com seu projeto de investigação, sua problemática e questão de pesquisa. De modo geral, dois aspectos sempre são considerados: os autores dos livros e suas atuações no momento histórico e a relevância das editoras, a partir, inclusive, de estudos da área de História da Educação sobre casas editoriais e personagens de destaque nas reformas educacionais. Por fim, temos, no momento, cerca de 150 livros digitalizados e organizados em uma planilha compartilhada no drive do GEPGE.

Justamente por ser um desafio e uma grande responsabilidade a criação de fontes digitais – no caso, livros didáticos digitalizados a serem mobilizados como fontes de pesquisa para a produção de Histórias da educação matemática – consideramos relevante compartilhar nossa experiência. Relatamos o processo que desenvolvemos, os cuidados, os erros e deslizes enfrentados, bem como as descobertas realizadas durante as práticas de digitalização conduzidas por um grupo coletivo.

É importante destacar que nosso objetivo não foi a constituição de um acervo digital – ainda que muitos dos livros digitalizados tenham sido compartilhados no Repositório Digital da UFSC –, mas sim o desenvolvimento de práticas de digitalização com recursos acessíveis e de baixo custo. Essa escolha visa atender uma demanda comum entre pesquisadores brasileiros

⁸ Metadados referem-se a dados sobre os dados, ou seja, informações adicionais que descrevem o arquivo. Em um arquivo digitalizado, eles podem incluir detalhes como a data da digitalização, o nome do digitalizador, o tipo de equipamento usado, as configurações de digitalização (resolução, formato), e até mesmo informações contextuais sobre o documento original (título, autor, data de criação).

que, em geral, não dispõem de infraestrutura tecnológica sofisticada ou equipamentos específicos.

A sistematização de um processo de digitalização acessível e replicável tende a modificar a forma como os pesquisadores lidam com as fontes históricas, permitindo a ampliação dos acervos disponíveis. Ao favorecer o compartilhamento de procedimentos técnicos e de bases digitais comuns, a proposta incentiva a colaboração entre instituições e grupos de pesquisa, contribuindo para a consolidação de uma cultura de cooperação e de preservação da memória educacional. Essa perspectiva aponta para um cenário no qual o trabalho historiográfico se torna mais coletivo, transparente e aberto à reinterpretação contínua das fontes.

Apresentamos, primeiramente, os espaços em que os livros didáticos físicos estavam armazenados – alguns no Centro de Documentação da Memória Científica e Pedagógica do Ensino da Matemática (CEMAT) e outros na Biblioteca do Livro Didático (BLD). Em seguida, descrevemos nosso processo de digitalização, acompanhado de comentários e sugestões que acreditamos serem úteis a pesquisadores da área de História, bem como de outras áreas que igualmente demandam a digitalização de documentos.

2 Acervos consultados

Antes de descrevermos os acervos visitados, é importante destacar os critérios que orientaram sua escolha e o papel que esses espaços desempenham nas práticas de pesquisa em História da educação matemática. A seleção do CEMAT e da BLD deve-se à sua relevância histórica, à amplitude dos materiais disponíveis e à representatividade das obras em relação ao ensino de matemática no Brasil. Ambos os acervos configuram-se como espaços de preservação e produção de conhecimento, valorizando a memória da educação. Os dois acervos consultados (CEMAT e BLD) estão localizados em cidades de grande circulação e de fácil acesso, o que favorece as atividades de pesquisa e visitação, como veremos a seguir.

2.1 CEMAT

O CEMAT é um espaço dedicado à preservação e ao estudo de documentos históricos relacionados ao ensino da matemática no Brasil. Sua criação está vinculada ao GHEMAT-SP, fundado em 2000 na PUC-SP, por pesquisadores interessados em compreender a trajetória da matemática enquanto disciplina escolar e suas práticas de ensino ao longo do tempo.

De acordo com a história relatada em seu site⁹, o centro de documentação foi formalizado em 2008, no contexto das ações do GHEMAT-SP, com a instalação de um acervo físico na cidade de Osasco (SP). Desde então, o CEMAT se consolidou como um importante polo de pesquisa, reunindo materiais como livros didáticos, periódicos, programas curriculares e documentos institucionais – fontes fundamentais para a reconstituição de práticas educativas e para a análise das políticas e métodos pedagógicos utilizados no ensino da matemática.

Um dos diferenciais do CEMAT é a reunião de acervos pessoais de educadores, pesquisadores e autores que desempenharam papéis importantes na História da educação matemática brasileira. Destaca-se, entre os acervos¹⁰ organizados ou em processo de organização, o Arquivo Pessoal de Ubiratan D'Ambrosio (APUA), que contém aproximadamente 500 caixas de arquivos com livros, artigos, correspondências, fotografias, rascunhos de obras, discursos e documentos de sua atuação nacional e internacional desde os anos 1970.

Além da preservação documental, o CEMAT atua como espaço de produção e difusão do conhecimento. Realiza seminários, encontros formativos e eventos acadêmicos. Atualmente, o centro funciona mediante agendamento¹¹ prévio para consultas presenciais, oferecendo suporte técnico e biblioteconômico especializado. Ao preservar e disponibilizar documentos essenciais para a História da educação matemática, o CEMAT contribui para a valorização da memória educativa como patrimônio histórico e fonte de pesquisa.

Nos últimos anos, o CEMAT passou por uma reestruturação significativa, motivada pelo crescimento expressivo de seu acervo e pela demanda por melhores condições de conservação e acesso aos documentos. Diante disso, o GHEMAT-Brasil adquiriu um espaço próprio e definitivo para sediar o centro, conferindo-lhe autonomia e melhores condições institucionais. Essa nova sede passou a abrigar formalmente o CEMAT, sendo agora mantido diretamente pelo GHEMAT-Brasil. A nova instalação está localizada na cidade de Santos, no litoral do estado de São Paulo.

⁹ Disponível em: <https://www.cemat-ghemat.com/sobre>. Acesso em: 27 jun. 2025.

¹⁰ Alguns outros acervos pessoais encontrados no CEMAT: Anna Franchi, Euclides de Medeiros Guimarães Roxo, Jacy Monteiro, Lucia Maria Aversa Villela, Lucília Bechara Sanchez, Luiz Márcio Pereira Imenes, Lourdes de la Rosa Onuchic, Lydia Condé Lamparelli, Manhúcia Liberman, Manoel Jairo Bezerra, Maria Amabile Mansutti, Maria do Carmo Domite, Neuza Bertoni Pinto, Osvaldo Sangiorgi, Scipione Di Pierro Netto e Sônia Coelho.

¹¹ Os agendamentos devem ser realizados por meio do site: <https://www.cemat-ghemat.com/sobre>. O funcionamento do centro é de segunda a quinta-feira, das 8h às 20h, e aos sábados, das 9h às 19h. Localizado na rua Carvalho Mendonça, 93 – conjunto 32, Vila Belmiro, Santos/SP – CEP 11079-100.

2.2 BLD

A Biblioteca do Livro Didático¹² está vinculada à Faculdade de Educação da Universidade de São Paulo (FEUSP), de acordo com o seu site¹³, configura-se como um acervo singular no Brasil em termos de quantidade, diversidade e antiguidade na área da Educação. A coleção abrange obras desde o século XIX até os dias atuais e constitui um rico campo de investigação para diferentes abordagens. Segundo o site da BLD, o livro didático, em sua multiplicidade de significados, pode ser analisado como artefato cultural, produto inserido no mercado editorial, meio de sistematização de saberes escolares ou, ainda, como transmissor de valores, ideologias e culturas.

A história do acervo da BLD teve início a partir dos anos finais dos anos 1980, consolidando-se entre 1994 e 1998, impulsionada por doações de pesquisadores que tinham o livro didático como foco de suas investigações ou o empregavam em processos formativos voltados à prática docente. Parte importante dessa iniciativa buscava suprir uma lacuna enfrentada por estudiosos da área: a escassez de livros didáticos disponíveis em bibliotecas, especialmente aqueles destinados ao ensino fundamental e médio, frequentemente descartados ou vistos apenas como instrumentos de reprodução ideológica (Bittencourt; Silva; Boto, 2021).

As pesquisas realizadas na FEUSP revelaram a natureza distinta dos livros didáticos em relação a outros impressos como a sua produção em grande escala, o vínculo estreito com as diretrizes curriculares e sua rápida obsolescência, conferem a esses materiais uma vida útil limitada. Justamente por essa característica transitória, os livros escolares raramente eram preservados por bibliotecas convencionais, o que justificou a criação de um espaço específico para sua guarda e estudo (Bittencourt; Silva; Boto, 2021). A organização da BLD envolveu não apenas a coleta sistemática dos materiais, mas também o desenvolvimento de práticas específicas de armazenamento, classificação e disponibilização voltadas a atender a comunidade acadêmica e pesquisadores interessados no tema.

A fase inicial da BLD esteve articulada ao subprojeto *Literatura escolar e memória escolar brasileira*, sob coordenação da professora Circe Bittencourt, vinculado ao Centro de Memória da FEUSP e inserido no projeto temático *Impressos, Leituras e Instituições Escolares no Brasil*, coordenado por Marta Chagas de Carvalho. Ao longo do tempo, a biblioteca se expandiu graças a doações feitas por professores, editoras e livreiros, transformando-se em uma das coleções especiais da Biblioteca da Faculdade de Educação. Além dos livros didáticos

¹² Os agendamentos para visitas são feitos pelo e-mail bldfe@usp.br. A biblioteca está localizada na Av. da Universidade, 308 - Butantã, São Paulo - SP, 05508-040; com horário de funcionamento de segunda a sexta-feira, das 13h às 18h.

¹³ Disponível em: <https://www4.fe.usp.br/biblioteca/acervos/biblioteca-do-livro-didatico>. Acesso em 27 jun. 2025.

convencionais, o acervo passou a incluir obras produzidas para e por escolas indígenas, assim como materiais voltados à Educação de Jovens e Adultos, cuja lógica de produção, circulação e uso é distinta da dos livros escolares tradicionais. De forma ampla, a BLD adota o critério de reunir todas as obras que participam do processo de aprendizagem, seja por meio de uma progressão sistemática de conteúdos ou como instrumentos de apoio, como dicionários, coletâneas literárias e materiais de consulta.

Um diferencial importante da BLD é o valor atribuído aos exemplares usados, que carregam marcas das práticas escolares. Rabiscos, anotações, respostas e dedicatórias tornam-se objetos de estudo ao revelarem o uso cotidiano do livro em sala de aula, dialogando diretamente com múltiplas possibilidades de análise, como a história da formação docente, das disciplinas escolares e dos métodos de ensino.

De acordo com Bittencourt, Silva e Boto (2021), com o avanço das pesquisas sobre a relação entre Estado, mercado editorial e políticas públicas, o acervo da BLD passou a incluir materiais relacionados ao Programa Nacional do Livro Didático (PNLD), além de legislações, guias do professor, catálogos de editoras e produções acadêmicas. Essa ampliação demandou uma categorização própria, orientada pelas particularidades da produção didática, para facilitar o acesso e a pesquisa qualificada.

Desde sua criação, a BLD esteve conectada a um Banco de Dados com o objetivo de mapear e localizar livros didáticos brasileiros em diversas bibliotecas do país e do exterior. Esse projeto resultou no desenvolvimento do Banco LIVRES (Livros Escolares), com apoio da Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), em articulação com a plataforma francesa Emmanuelle e com instituições como a Bibliothèque Nationale de France. A adaptação desse modelo à realidade brasileira permitiu destacar elementos específicos da circulação e da edição escolar no país.

A partir de 2016, o LIVRES passou a integrar o projeto *International Textbook Catalogue*, coordenado pelo *Georg Eckert Institute for International Textbook Research*, da Alemanha, em parceria com acervos da Espanha (MANES), Itália (EDISCO) e Brasil. A rede promove a digitalização, a padronização técnica e a circulação global de livros escolares, dentro dos princípios das Humanidades Digitais. No âmbito da USP, esse esforço se consolidou na criação do Global Textbook Resource Center (GLOTREC), que atua na integração de bases de dados e desenvolvimento de soluções técnicas de biblioteconomia e informática.

Atualmente, a BLD conta com aproximadamente 23 mil livros e 1.200 documentos, e está inserida no projeto temático *Arquivos digitais e bibliotecas: história do livro e da leitura*, um dos eixos da iniciativa *FAPESP Saberes em Fronteira: por uma história transnacional da*

educação (1810). Por meio desse conjunto de ações, a biblioteca se consolida como espaço de preservação da memória escolar brasileira e de promoção de investigações fundamentais sobre o papel do livro didático nos processos educativos do passado e do presente.

3 Práticas de digitalização

A possibilidade de digitalização de materiais impressos não é recente: os primeiros *scanners* comerciais e domésticos foram lançados na década de 1980. Desde então, os equipamentos evoluíram consideravelmente, chegando aos atuais celulares equipados com câmeras fotográficas de alta resolução, o que ampliou as possibilidades de digitalização de livros e documentos.

Durante muito tempo, o processo de digitalização, para a maioria dos pesquisadores, consistia basicamente em posicionar o documento sobre a mesa do *scanner* e acioná-lo, normalmente por meio de um comando no computador conectado ao equipamento. Essa dinâmica, quando realizada por uma única pessoa, exigia inúmeras repetições do procedimento: retirar o livro da mesa, virar a página, reposicioná-lo, acionar o *scanner* – eventualmente percorrendo o trajeto entre o computador e o *scanner* a cada etapa.

Um exemplo interessante de práticas de organização, digitalização e constituição de acervos digitais, desenvolvidas por historiadores da educação matemática, pode ser encontrado no capítulo *Para guardar o que quer que se guarde: dos acervos escolares à construção de uma coleção digital*, de Rios e Rodrigues (2020). Nesse relato, descreve-se o processo de digitalização de documentos com *scanners* de alta qualidade realizado por pesquisadores – bolsistas de Iniciação Científica, mestrandos e professores da Universidade Federal do Rio Grande do Sul/UFRGS, Universidade Federal de Pelotas/UFPel, Universidade de Passo Fundo/UPF) – no âmbito do Projeto de Pesquisa *Estudar para ensinar: práticas e saberes matemáticos nas escolas normais do Rio Grande do Sul (1889-1970)*, financiado pelo CNPq. A Figura 1, apresentada a seguir, foi obtida a partir do acervo do referido projeto.

Figura 1 – Digitalização de documentos em *scanners* de alta qualidade



Fonte: Rios e Rodrigues (2020, p. 80)

Com o advento de celulares equipados com câmeras fotográficas, tornou-se possível realizar a digitalização rápida de páginas de livros e outros documentos, evitando, por exemplo, a necessidade de empréstimo de obras em bibliotecas ou múltiplas visitas a acervos com exemplares raros, que não circulam.

Nossa experiência com a digitalização de documentos – especialmente livros didáticos – teve início em 2023, período em que celulares equipados com câmeras de alta resolução já estavam amplamente disponíveis. Para aumentar a produtividade na captura de centenas de páginas, optamos por nos inspirar na estrutura dos chamados *scanners* planetários, equipamentos de alto custo compostos por três elementos principais: uma base (geralmente na cor preta para realçar o contraste com a cor das páginas), um suporte elevado para o sistema óptico (a câmera) e um sistema de iluminação, geralmente com luzes de LED laterais ou anguladas, que proporcionam iluminação uniforme, baixa emissão de calor, evitam reflexos excessivos e não danificam o papel.

Nossa primeira tentativa de simular essa estrutura envolveu o uso de um tripé para manter o celular fixo e de radiografias escuras como base. No entanto, desde os testes iniciais, realizados no CEMAT, a iluminação revelou-se um fator determinante para o sucesso na captura automática. As tentativas de utilizar luz ambiente, abajures ou lanternas resultaram, na maioria das vezes, em iluminação irregular e sombras, comprometendo a qualidade das imagens.

É importante destacar que as fontes digitais que produzimos têm como principal objetivo servir de material de consulta para o nosso grupo de pesquisa. Nesse contexto, o conteúdo textual e as imagens constituem o objeto de análise e, por essa razão, diferentemente

dos Arquivos Digitais discutidos por Santos (2021), não buscamos gerar arquivos com qualidade editorial, mas sim disponibilizar fontes de consulta fidedignas aos livros originais, que possibilitem a continuidade das pesquisas mesmo à distância de centros como o CEMAT e a BLD. Considerando essa finalidade, aceitamos algumas imperfeições, como pequenas sombras sobre as páginas e variações de cor entre o original físico e a versão digitalizada, desde que não comprometa a leitura e a análise dos conteúdos.

Nossa primeira estrutura para digitalização, ainda que limitada em termos de iluminação, possibilitou, já no primeiro ano de atividades, a organização de algumas coleções de obras de interesse para o grupo, digitalizadas no CEMAT, incluindo livros didáticos de autores como Euclides Roxo, Cecil Thiré e Mello e Souza e de Osvaldo Sangiorgi.

Contudo, apesar da portabilidade proporcionada pelo tripé, a inconsistência da iluminação nos diversos locais visitados nos levou a buscar uma alternativa mais eficiente. Nesse processo, encontramos um dispositivo dobrável multifuncional com iluminação embutida¹⁴. A Figura 2 ilustra essa nova estrutura: à esquerda, observa-se o livro didático, o celular e o dispositivo dobrável; à direita, o equipamento encontra-se aberto com o celular acoplado e o livro posicionado para ser fotografado.

Figura 2 – Dispositivo dobrável multifuncional (fechado e aberto)



Fonte: produzido pelos autores

O dispositivo adotado – com peso de aproximadamente 1,2 kg e custo relativamente baixo¹⁵ – mostrou-se ideal para o nosso processo de digitalização. Embora a iluminação fornecida não seja particularmente intensa ou completamente uniforme, o equipamento viabilizou digitalizações mais consistentes, sobretudo pela facilidade de ajuste da distância

¹⁴ Ressalta-se a contribuição do primeiro autor desse artigo, cujo conhecimento técnico foi imprescindível para orientar o grupo de pesquisa durante o desenvolvimento das digitalizações.

¹⁵ O dispositivo que usamos custava cerca de R\$ 220,00 em junho de 2025 em sites como Mercado Livre (<https://www.mercadolivre.com.br>).

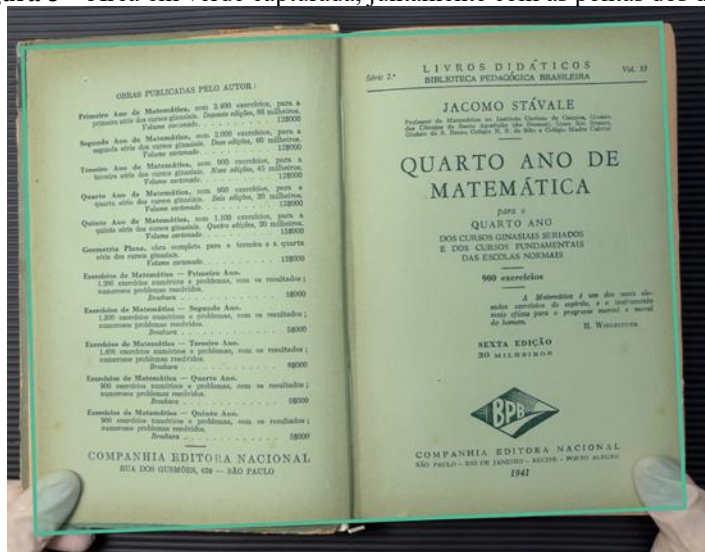
entre o documento e o celular (Figura 2, à direita). Além disso, por ser dobrável, pode ser facilmente transportado, o que é essencialmente útil para atividades em diferentes locais. Certamente pode haver dispositivos mais eficientes, mas reiteramos que ao encontrarmos o suporte da Figura 2, consideramos muito adequado em qualidade, custo e praticidade de transportar, atendendo aos nossos objetivos, e assim sendo, não realizamos estudos mais aprofundados de outros suportes.

A etapa fundamental nesse processo é a captura (digitalização) do documento e a posterior edição das imagens. À medida que as câmeras dos celulares passaram a oferecer altas resoluções, surgiram aplicativos específicos para digitalização, que incorporaram funcionalidades como reconhecimento de bordas e captura automática.

A escolha de um aplicativo para a digitalização deve considerar as funcionalidades disponíveis¹⁶. Praticamente todos os aplicativos atuais oferecem as duas funções que consideramos essenciais para a produtividade: reconhecimento automático de bordas e captura sem necessidade de acionamento manual. Essas funcionalidades representam um ganho significativo de produtividade em relação ao processo de digitalização com *scanner* de mesa.

O reconhecimento de borda (em verde na Figura 3) elimina, durante a edição do arquivo gerado, a necessidade de recorte manual das páginas. Esse reconhecimento, dependendo do aplicativo utilizado, ocorre de forma quase instantânea e mantém a eficácia mesmo quando o enquadramento apresenta inclinação.

Figura 3 – Área em verde capturada, juntamente com as pontas dos dedos



Fonte: produzido pelos autores

¹⁶ Testamos aplicativos como Scanner Pro (da Readdle), Lens (da Microsoft), Google Drive e CamScanner (<https://www.camscanner.com>), sendo este último o que consideramos mais eficiente, especialmente quanto à velocidade de reconhecimento das bordas e à captura. Em nossa experiência, todos se mostraram similares e responderam bem ao propósito dos ajustes que realizamos, sendo que aqueles com versões pagas cobravam, em julho de 2025, a partir de R\$ 16,90 por mês.

A funcionalidade da captura automática, combinada com o reconhecimento de bordas, permite digitalizar um livro inteiro *apenas* folheando suas páginas. Após abrir o exemplar na primeira página a ser digitalizada, o aplicativo identifica automaticamente os contornos da página e realiza a captura após alguns instantes. Em seguida, basta virar a próxima página, e o processo se repete, dispensando acionamento manual.

Com algum tempo de familiarização com o aplicativo, esse procedimento pode se tornar extremamente ágil: é possível digitalizar um livro de 300 páginas em cerca de 10 minutos, considerando a captura de duas páginas por imagem. A Figura 4 ilustra uma das pesquisadoras do grupo realizando esse processo de digitalização na BLD.

Figura 4 – Detalhes do trabalho de digitalização na BLD



Fonte: acervo do GEPGE

A digitalização resulta em um arquivo PDF, formato que mantém fidelidade ao documento original – preservando layout e conteúdo – independentemente do sistema operacional ou dispositivo em que é visualizado. A etapa final – a edição do arquivo PDF – inclui procedimentos como a *retirada dos dedos*, a *separação das páginas* e a chamada *ocerização* (neologismo informal derivado do inglês OCRred ou OCR conversion, sendo OCR a sigla para Optical Character Recognition, ou reconhecimento óptico de caracteres).

Alguns aplicativos de digitalização oferecem funções auxiliares que podem ser aplicadas já durante a captura ou posteriormente, tais como: remoção de marcas de dedo (que será discutida adiante), clareamento das páginas, eliminação de anotações manuscritas e conversão para a escala de cinza. Essas funções, comumente chamadas de filtros, devem ser testadas previamente, especialmente no caso de documentos extensos, para evitar distorções ou perdas indesejadas.

Por exemplo, o filtro de clareamento de páginas é útil para eliminar o efeito de transparência do verso da folha (frequente em papéis finos) e melhorar a eficácia na remoção de marcas de dedo. No entanto, pode comprometer a leitura de elementos em vermelho ou amarelo, tornando-os ilegíveis. Da mesma forma, o filtro de remoção de escrita à mão pode ser bastante eficaz quando o objetivo é preservar apenas o conteúdo do impresso, mas sua aplicação deve ser cuidadosamente avaliada caso as anotações manuscritas sejam relevantes para a pesquisa¹⁷.

Portanto, a escolha e a aplicação de um filtro devem ser definidas em consonância com os objetivos da digitalização e testadas previamente, a fim de evitar a necessidade de refazer o processo – o que pode ser inviável, especialmente quando o material original está localizado a grande distância do pesquisador.

Em nossa experiência, embora os aplicativos de captura ofereçam funções de edição, observamos melhores resultados quando a captura é feita pelo celular e o processamento posterior é realizado em um computador. A utilização de softwares específicos para edição de PDF¹⁸, combinada à visualização em tela maior, oferece maior precisão, especialmente o tratamento de elementos gráficos.

Assim como ocorre nos aplicativos de captura, há uma grande variedade de opções disponíveis para a edição dos arquivos PDF gerados na digitalização. A maioria desses softwares compartilha funcionalidades semelhantes, voltadas à melhoria do arquivo final para fins de leitura, análise e pesquisa.

Similarmente ao funcionamento dos *scanners* de mesa – nos quais é comum a captura de uma imagem que reúne as duas páginas abertas de um livro – os arquivos obtidos em nossos procedimentos frequentemente seguem esse padrão. Por esta razão, a etapa de edição geralmente envolve as seguintes etapas:

- a) *exclusão de páginas duplicadas*, com conferência cuidadosa para garantir que todas as páginas tenham sido devidamente capturadas;
- b) *correção de distorções*, especialmente em imagens com inclinação acentuada ou desalinhadas;

¹⁷ Apagar marcações feitas por leitores é uma decisão a ser tomada com prudência, pois além de informações históricas complementares que algumas notas podem trazer, algumas vezes é possível inferir data limite da publicação, para obras que não tenham essa informação.

¹⁸ Testamos os aplicativos iLovePDF, Smallpdf, PDF2Go, SwifDoo PDF e PDF Expert (<https://pdfexpert.com/pt>), sendo que este último se mostrou muito superior aos demais, porém só é compatível com os sistemas operacionais da Apple. Em julho de 2025, a mensalidade desses aplicativos partia de R\$ 20,00. Em nossa experiência, todos os aplicativos testados foram capazes de executar cada uma das etapas do processamento pós-captura.

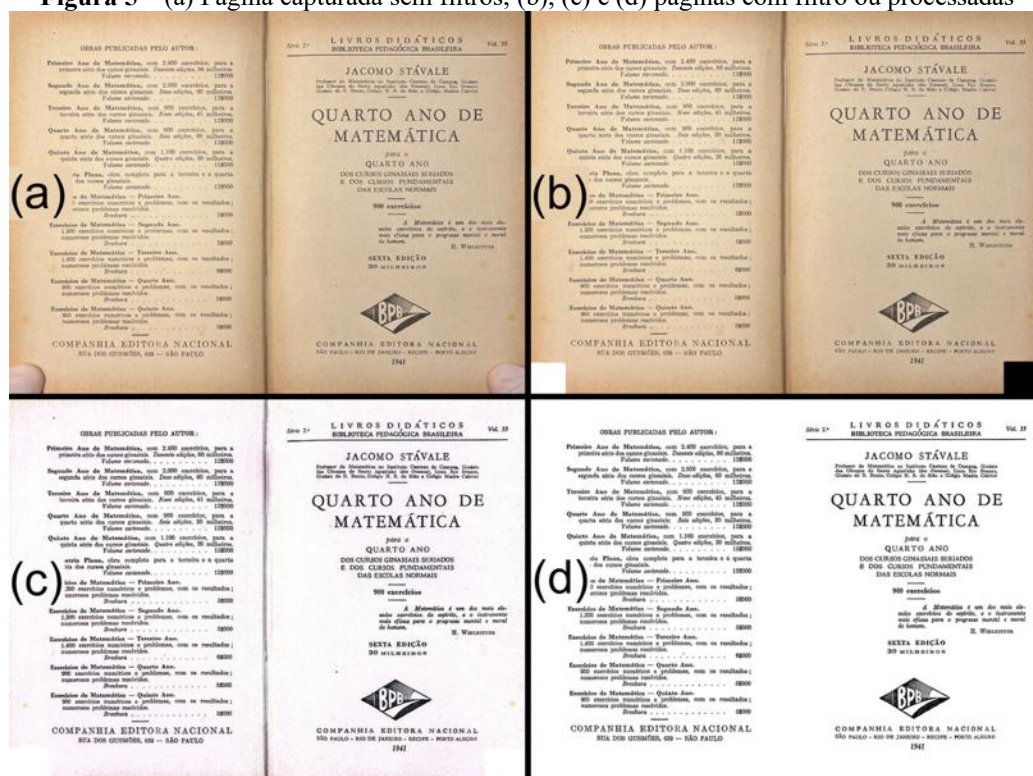
- c) *remoção de marcas de dedos* e anotações manuscritas nas margens, quando não pertinentes à pesquisa;
- d) *separação das páginas capturadas em conjunto*, de modo que cada página componha uma unidade individual do PDF;
- e) *uniformização do tamanho das páginas*, preferencialmente tomando como referência as dimensões do exemplar original;
- f) *aplicação do reconhecimento óptico de caracteres (OCR)*, etapa que torna o arquivo *pesquisável*, permitindo localizar palavras-chaves, contar ocorrências e realizar outras análises textuais diretamente no documento digitalizado.

O tempo de edição do arquivo PDF digitalizado depende do propósito da pesquisa. Caso a intenção seja apenas obter um texto pesquisável e com páginas uniformes — o que consideramos favorecer uma leitura fluida —, as etapas (a), (d), (e) e (f) são relativamente rápidas e podem ser todas executadas em cerca de 10 minutos.

Se o arquivo a ser gerado for compartilhado com outros pesquisadores por meio de repositórios, entendemos ser fundamental que a etapa (c) seja executada. A depender da fonte, essa é a etapa mais demorada. Arquivos com folhas soltas ou encadernados em espiral podem ser capturados sem marcas de dedos, o que elimina a necessidade dessa etapa; porém, livros, em geral, exigem o uso das mãos para se manterem abertos — por vezes, em todas as páginas. Nesse último caso, apenas a etapa (c) pode levar mais de 30 minutos para a remoção das marcas de dedos em 300 páginas.

A Figura 5(a) ilustra o tipo de página comumente obtida no processo de captura. Como muitos livros exigem o uso das mãos para se manterem abertos, as pontas dos dedos dos pesquisadores frequentemente aparecem nas imagens — o que é um elemento visualmente incômodo, pois pode chamar demais a atenção no PDF gerado. Por esse motivo, uma das etapas mais trabalhosas durante a edição costuma ser a *remoção das marcas de dedos*, que, em geral, deve ser feita manualmente, página por página. A Figura 5 apresenta um exemplo de página originalmente capturada, seguida de três alternativas de edição, evidenciando estratégias diferentes de remoção das marcas.

Figura 5 – (a) Página capturada sem filtros; (b), (c) e (d) páginas com filtro ou processadas



Fonte: produzido pelos autores

Observe na Figura 5(a) que a captura original preserva as cores da página, de acordo com o que o celular *enxerga*, que inclui as pontas dos dedos de quem digitalizou. A Figura 5(b) mostra uma opção comum para retirar os dedos, que consiste em censurar¹⁹ a área que não desejamos que apareça, normalmente com retângulos brancos ou pretos. Embora o efeito de remoção seja alcançado, isso deixa a impressão de página recortada.

Na Figura 5(c) há uma página que passou por um filtro de clareamento que não removeu os dedos, porém, após a *censura manual das marcas*, a página se tornou mais agradável visualmente, com a área removida bem mais discreta. Finalmente, a Figura 5(d) mostra uma página que passou por um filtro de clareamento mais avançado, o que removeu automaticamente as marcas dos dedos.

Esse último filtro deve ser analisado conforme os objetivos da digitalização, pois em textos com muitas figuras ou cores, é possível que parte do conteúdo também seja excluído. Por essa razão, reiteramos a recomendação de testar os filtros antes de iniciar a digitalização de um documento. Caso a fidelidade às cores e à definição, tal como se apresentam na fonte original, seja o aspecto mais importante para os propósitos da pesquisa, a captura sem filtros é a que

¹⁹ Em alguns aplicativos essa opção é chamada *redact* (redigir), cuja função é remover permanentemente informações sensíveis do PDF. No entanto, na maioria dos aplicativos que testamos, apesar de conseguirmos o efeito desejado de apagar os dedos, ao passar o PDF pelo processo de OCR no aplicativo PDF Expert, os dedos *reapareceram*, o que indica uma falha na garantia desses aplicativos de que áreas censuradas não podem ser recuperadas.

proporciona a melhor cópia. As figuras 5(a) e 5(b) são muito próximas das páginas originais no momento da captura — ou seja, com as páginas escurecidas pelo efeito do tempo. Ironicamente, a figura 5(d) pode representar a página como era ao sair da gráfica, branca; mas, como não podemos fazer essa afirmação, mantemos o princípio de que cada cópia que realizamos

exige uma forma relativamente fixa, em termos de um conteúdo estável, uma procedência e um contexto que assegurem que a digitalização foi bem-sucedida: a inteireza do conteúdo necessita *de fato* estar presente na cópia digitalizada, aquilo que os arquivistas denominam cadeia de custódia. Em outros termos, qualquer tipo de erro, negligência ou até mesmo má-fé no processo computacional pertinente à digitalização será determinante no trabalho historiográfico (Brasil; Nascimento, 2020, p. 201, grifo dos autores).

Partindo do princípio de que as fotos tiradas pelos aplicativos de captura e tratadas pelos aplicativos de edição mantêm o conteúdo praticamente inalterado, entendemos que as fontes digitais que produzimos são *verdadeiras*, idênticas em conteúdo às fontes originais e seguras para serem compartilhadas com a comunidade de pesquisadores.

4 Considerações finais

Em concordância com Michel De Certeau, com quem iniciamos este artigo, a prática que compartilhamos representa a primeira etapa do fazer historiográfico: a identificação de acervos de interesse e a captura dos documentos a serem analisados. Embora compreendamos que a disponibilização de fontes digitais — especialmente aquelas oriundas de processos de digitalização conduzidos por Arquivos Digitais, como os da Biblioteca Nacional Digital e do Repositório de Conteúdo Digital da Universidade Federal de Santa Catarina — tende a se tornar cada vez mais abundante, também reconhecemos que muitos pesquisadores ainda dependerão de seu próprio trabalho para separar, reunir e transformar em documentos *digitais* inúmeros vestígios do passado ainda não tratados pelos Arquivos institucionais.

Os processos que descrevemos corroboram a posição de Brasil e Nascimento (2020, p. 203), segundo os quais “o trabalho do historiador diante do arquivo digital, portanto, não é tão diferente do trabalho diante do arquivo físico, pois exige tanto rigor metodológico no tratamento da fonte quanto o tratamento de uma fonte não digital”. Em nossa prática, buscamos a maior fidelidade possível às fontes a partir de uma estrutura de captura não profissional de maneira a permitir que os pesquisadores distantes dessas fontes possam analisá-las com segurança.

Destacamos que os custos envolvidos são relativamente baixos, pois partimos do princípio de que todo pesquisador já dispõe de um aparelho celular com câmera de alta resolução, sendo necessário apenas um suporte adequado para o início das digitalizações. Os

aplicativos para captura e edição constituem os recursos complementares e, embora seja possível utilizá-los gratuitamente, algumas das melhores funcionalidades só estão disponíveis em versões pagas. Além disso, ao utilizar um aplicativo gratuito, é comum a presença de marcas d'água com o nome das empresas desenvolvedoras em cada página digitalizada ou processada, o que gera quase o mesmo incômodo causado pela aparição das *pontas dos dedos*.

Ainda que o GEPGE não tenha a pretensão de se constituir como um repositório digital, alguns dos materiais produzidos já foram compartilhados com o Repositório Digital Institucional da UFSC, colaborando para a ampliação do acesso e a preservação de obras. Acreditamos que essas contribuições podem fomentar parcerias e auxiliar no processo de democratização dos acervos.

Por fim, consideramos essencial reafirmar a importância da confiabilidade das fontes digitais, especialmente no que se refere à preservação dos metadados e à rastreabilidade do processo de digitalização. Como afirmam Brasil e Nascimento (2020), são os metadados que explicam, contextualizam e conferem veracidade ao documento. Nesse sentido, o cuidado com a descrição, a procedência e as condições de produção das cópias digitais é parte integrante do trabalho historiográfico, podendo inclusive envolver algum tipo de certificação da autoria do processo – como assinaturas digitais ou selos de autenticação. É somente a partir desse esforço coletivo de registro, rigor técnico e partilha que poderemos ampliar, com responsabilidade, os horizontes da pesquisa em meio digital.

Além disso, realizamos a validação das cópias digitais por meio da comparação direta com os exemplares físicos, assegurando a correspondência integral das páginas, a legibilidade e a fidedignidade do conteúdo. Essa etapa foi conduzida de forma criteriosa por diferentes membros do grupo, de modo a garantir que as versões digitais preservassem o máximo possível as características dos originais e pudessem ser reconhecidas como fontes confiáveis para futuras pesquisas.

Agradecimentos

Nossa sincera gratidão à professora Ana Paula Jahn pela leitura prévia cuidadosa e pelos valiosos comentários ao nosso manuscrito. Sua contribuição foi fundamental para o aprimoramento deste trabalho.

Referências

- BITTENCOURT, Circe; SILVA, Carolina Mostaro Neves da; BOTO, Carlota. A Faculdade de Educação pela clivagem dos livros: Biblioteca do Livro Didático e Biblioteca Paulo Bourroul. **RIDPHE_R Revista Iberoamericana do Patrimônio Histórico-Educativo**, v. 7, p. 1–15, 2021. Disponível em: <https://econtents.bc.unicamp.br/inpec/index.php/ridphe/article/view/16060>. Acesso em: 27 jun. 2025.
- BRASIL, Eric; NASCIMENTO, Leonardo Fernandes. História Digital: Reflexões a Partir da Hemeroteca Digital Brasileira e do uso de CAQDAS na Reelaboração da Pesquisa Histórica. **Estudos Históricos (Rio de Janeiro)**, v. 33, n. 69, p. 196–219, abr. 2020. Disponível em: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0103-21862020000100196&tlng=pt. Acesso em: 3 jun. 2025.
- DE CERTEAU, Michel. **A Escrita da História**. 2. ed. Tradução: Maria de Lourdes Menezes. Rio de Janeiro: Forense Universitária, 2002.
- GREGORIO, Janine Marques da Costa; COSTA, David Antonio da; MACIEL, Viviane Barros. Práticas historiográficas apoiadas nos objetos digitais: contribuições da História digital. **Revista Brasileira de História da Educação**, v. 25, n. 1, p. 1–25, 2025. Disponível em: <https://periodicos.uem.br/ojs/index.php/rbhe/article/view/74590>. Acesso em: 11 jun. 2025.
- RIOS, Diogo Franco; RODRIGUES, Janine Moscarelli. Para guardar o que quer que se guarde: dos acervos escolares à construção de uma coleção digital. In: BÚRIGO, Elisabete Zardo *et al.* (org.). **Saberes Matemáticos nas Escolas Normais do Rio Grande do Sul (1889-1970)**. São Leopoldo: Oikos, 2020. p. 69–90. *E-book*.
- SANTOS, Daise Silva dos. Arquivos Digitais: possibilidades de pesquisa no campo da História da Educação. **Caminhos da Educação: diálogos culturas e diversidades**, v. 3, n. 2, p. 4–19, 22 dez. 2021. Disponível em: <https://periodicos.ufpi.br/index.php/cedsd/article/view/2369>. Acesso em: 14 mai. 2025.
- VALENTE, Wagner Rodrigues. Oito temas sobre História da Educação Matemática. **REMATEC**, v. 8, n. 12, p. 22–50, 2013. Disponível em: <https://www.rematec.net.br/index.php/rematec/issue/view/32/16>. Acesso em: 4 jul. 2025.