

Entre papéis e algoritmos: desafios metodológicos da História da Educação nas Humanidades Digitais

Between papers and algorithms: methodological challenges of the History of Education in Digital Humanities

Alessandro Santos da Rocha

Doutor em Educação pela Universidade Estadual de Maringá, Brasil
Professor do Departamento de Fundamentos da Educação (DFE) e do Programa de Pós-graduação em Educação (PPE) da Universidade Estadual de Maringá, Brasil
asrocha2@uem.br
<http://lattes.cnpq.br/1622339109413385>
<https://orcid.org/0000-0002-2664-5224>

Italo Ariel Zanelato

Doutor em Educação pela Universidade Estadual de Maringá, Brasil
Professor do Departamento de Fundamentos da Educação (DFE) da Universidade Estadual de Maringá, Brasil
iazanelato2@uem.br
<http://lattes.cnpq.br/6383304497323065>
<https://orcid.org/0000-0003-2072-8405>

Resumo: Este artigo analisa os desafios metodológicos da pesquisa com a imprensa periódica no campo da História da Educação, a partir das Humanidades Digitais. Historicamente, jornais e revistas constituíram espaços privilegiados de disputas e de projetos pedagógicos, nos quais intelectuais buscaram conformar a opinião pública em torno de determinadas hegemonias educativas. Se antes o trabalho investigativo implicava a crítica da materialidade da fonte, hoje ele se confronta com mediações técnicas, como o OCR e o CAQDAS, que interferem na seleção e na visibilidade dos documentos. Fundamentado no método de pesquisa bibliográfica com revisão da literatura, o texto demonstra que a pesquisa histórica precisa articular o potencial técnico das plataformas com rigor metodológico, a fim de preservar a densidade da pesquisa em História da Educação a partir das fontes da imprensa periódica.

Palavras-chaves: História da Educação; Imprensa Periódica; Humanidades Digitais.

Abstract: This article analyzes the methodological challenges involved in research on periodical press sources within the field of the History of Education, in light of Digital Humanities. Historically, newspapers and magazines constituted privileged spaces for disputes and pedagogical projects, in which intellectuals sought to shape public opinion around specific educational hegemonies. Whereas scholarly work previously required close attention to the materiality of the source, it now confronts technical mediations such as OCR and CAQDAS, which affect the selection and visibility of documents. Grounded in bibliographical research and a review of literature, the study demonstrates that historical inquiry must articulate the technical potential of digital platforms with methodological rigor in order to preserve the density of research in the History of Education based on periodical press sources.

Keywords: History of Education; Periodical Press; Digital Humanities.

Introdução

O trabalho do historiador da educação que se debruça sobre investigações com a imprensa periódica apresenta-se profundamente ligado à experiência concreta e à realidade de acervos, bibliotecas e arquivos. Em termos de execução prática, durante muito tempo, a pesquisa sobre os periódicos exigiu deslocamentos físicos, manuseio cuidadoso e conhecimento técnico para o trato com exemplares que nem sempre apresentavam condições para o trabalho corriqueiro sem técnicas específicas, haja vista tratar-se de documentos fragilizados por diversas questões, como o próprio decorrer do tempo ou as condições de armazenamento.

Evidentemente, a pesquisa em fontes impressas carrega consigo uma relação sensorial: o peso do papel, a tipografia, as marcas do uso e da circulação, ou seja, questões que integravam o processo técnico-metodológico. Assim, a imprensa não deve ser lida apenas como registro informativo no qual o pesquisador se debruça para buscar informações. Ela deve ser analisada em todas as suas potencialidades, principalmente em História da Educação, onde jornais e revistas são dispositivos pedagógicos por excelência. Notadamente, interessam ao historiador da educação características essenciais para a crítica de suas fontes, como a disposição das matérias, os métodos de ensino anunciados e as contendas que tratam sobre o fenômeno educativo, dentre outros. Não se pode perder de vista que o jornal atuava com objetivos que adentram os interesses pedagógicos, não sendo mero repositório de informações. Conceber a imprensa dessa forma é algo que deve permanecer, mesmo frente às leituras aceleradas feitas a partir de um periódico que antes exigia um tempo de leitura analógica e agora, por vezes, é consumido rapidamente em uma tela de computador.

Nas últimas décadas, entretanto, essa experiência foi sendo progressivamente transformada. A digitalização de acervos e a criação de hemerotecas virtuais alteraram não apenas as condições de acesso às fontes, mas também a própria dinâmica da investigação histórica. O que antes exigia semanas de busca, recorrendo a vários arquivos, pode, hoje, ser localizado em segundos por meio de palavras-chave dentro de buscadores cada vez mais dinâmicos em plataformas com múltiplos recursos. Essa mudança não se restringe a uma otimização técnica; ela redefine as mediações entre pesquisador e documento.

Importa lembrar que, para o campo educativo, a imprensa funcionava como um laboratório de projetos de reforma social. Por meio das folhas impressas, intelectuais e

reformadores disputavam a hegemonia sobre o que deveria ser a escola e como o sujeito deveria ser formado. Quando o historiador se debruça sobre essas fontes, busca capturar não apenas o texto, mas a pedagogia implícita na organização do impresso, elemento que corre o risco de ser invisibilizado quando reduzido ao resultado automatizado de um sistema de busca.

Nesse cenário, impõe-se uma questão metodológica que orienta este artigo: como preservar a densidade interpretativa do historiador ao conceber a produção da fonte, quando a mesma passa a ser mediada por sistemas de busca no contexto das Humanidades Digitais? Para responder ao problema apresentado, o texto tem como objetivo discutir a imprensa enquanto fonte privilegiada para o historiador da educação e, posteriormente, apresentar o contexto da pesquisa em história da educação a partir do âmbito que se encontra as Humanidades Digitais, pretende-se, com esse percurso, evidenciar questões técnicas das plataformas digitais mediados por algoritmos e, por fim, a validade da técnica historiográfica no âmbito das Humanidades Digitais. O trabalho é respaldado pelo método bibliográfico, perfazendo uma revisão da literatura e a análise de autores que ampliam a perspectiva da pesquisa na área da História da Educação. Ao fim, busca-se demonstrar que o trabalho investigativo deve guardar a perspectiva da crítica ao documento, mesmo considerando os novos modelos de acesso e consulta. Não se trata de rejeitar tais ferramentas, mas de compreender os limites que elas impõem à leitura crítica.

A Imprensa como fonte da História da Educação e as Plataformas Digitais

A imprensa periódica se coloca como uma das principais fontes para os historiadores da educação, sobretudo para o entendimento de momentos em que jornais e revistas ocupavam o espaço privilegiado de circulação de ideias e de formação da opinião pública (BARROS, 2012). Historicamente, a produção da imprensa foi se transformando a partir das tecnologias disponíveis. Quando considerada a historicidade da produção registrada pela imprensa, são verificáveis as mudanças utilizadas para imprimir a letra sobre o papel, de Gutenberg, na Alemanha de meados do século XV, às editoras que usam tecnologias de ponta no século XXI. Diversas alterações ocorreram e decorrem delas não só a materialidade do registro, que

primeiramente utilizou papéis e hoje paira em plataformas com recursos que facilitam desde a alta quantidade de armazenamento até a localização precisa das informações.

As mudanças ocorridas na produção da imprensa não estão apenas no âmbito da sua produção, mas residem também em questões que são necessárias de serem compreendidas, a exemplo de a quem ela serviu, quais os grupos que se utilizaram dela para buscar a hegemonia e, para além disso, os debates ideológicos que envolvem parcialidade, neutralidade e a própria formação do jornalismo. No caso da imprensa brasileira, podemos encontrar autores clássicos, como Nelson Werneck Sodré, publicado pela primeira vez em 1966, que caracterizam esse percurso histórico da imprensa, mapeando os problemas econômico-sociais e deixando em evidência que a imprensa periódica ocupava lugar especial como agente político dentro da sociedade do capital, o que nos remete à clássica frase do autor ao afirmar que: “[...] a história da imprensa é a própria história do desenvolvimento da sociedade capitalista” (SODRÉ, 1999: 7).

A história da imprensa brasileira não pode ser entendida como neutra, ao contrário, a imprensa periódica assumiu posicionamentos ideológicos que podem ser entendidos como servis a diversos interesses. Por sua vez, o historiador ao lidar com jornais e revistas, deve ter a compreensão de que a imprensa não reflete a realidade. Ela é uma instituição que carrega diversos interesses, inclusive auxiliou para moldar o capitalismo. Nessa perspectiva concordamos com Ribeiro (2015: 280), que afirma: ser a imprensa uma “força histórica ativa e uma das principais instituições que deram forma ao mundo capitalista”.

Ciente de que a imprensa se apresentou vinculada aos contextos históricos aos quais serviu, é crível que se compreenda que ela tem volatilidade e, inclusive em seus formatos, alterou-se ao longo da história, seja na perspectiva da produção material, ou nos interesses a que ela atende. Assim, não é possível compreendê-la distante dessas alterações. Alguns momentos são precisos para abranger essa transformação, como aqueles que se ergueram com a chegada da imprensa oficial no Brasil, com a vinda da família real no início do século XIX, ou ainda as alterações mais rígidas ocorridas durante o século XX. Sobre isso, Tânia Regina de Luca ressalta determinadas mudanças que se produziram no início dos 1900, impondo inovações tecnológicas. Nas palavras da autora: “As inovações não se limitaram às mudanças na estrutura de produção, organização, direção e financiamento, mas atingiram também o conteúdo dos jornais e sua ordenação interna, que começou a exigir gama variada de competências, fruto da divisão do trabalho e da especialização” (LUCA, 2008: 152).

O mapeamento da imprensa brasileira, guardada toda a historicidade de sua produção, não é mero repertório para os historiadores, que não se limitam à informação nela contida. A compreensão de sua produção, aliada às heurísticas que adentram no ofício do historiador, faz com que a imprensa apresente dados singulares para contar a história. No caso da História da Educação, tal entendimento resguarda a possibilidade de reencontrar projetos de ordenamento social; todavia, nem sempre o ofício de localizar e levantar as fontes foi automático, sobretudo quando aquele que irá sistematizar a história desconhece as práticas presentes na formação do historiador, como a exigência técnico-metodológica, a começar por localizar onde estava o periódico até a identificação das edições que poderiam compor seu quadro discursivo.

Notadamente, as questões acima apresentadas, desde a produção da imprensa periódica ao trabalho do historiador que a tem como fonte, alteraram-se. Da mesma forma, as perguntas também se alteraram, a exemplo: como proceder com o exercício de historiar a educação quando os documentos saem da esfera física para a digital? A resposta perpassa e centra-se na nova forma de encontrar os documentos, haja vista que: “A disponibilidade de um *corpus* documental de difícil acesso em formato digital possibilita promover novas possibilidades de investigação de temas que ainda se encontram obscuros na História da Educação brasileira” (NASCIMENTO; ALVES, 2019: 13).

Esse novo *corpus* documental, no caso específico da História da Educação, ganha contornos a serem repensados. Exemplo disso são os jornais da imprensa anarquista ou os jornais republicanos da virada do século XIX e XX que, na ausência de digitalização, ficaram restritos à folha de papel. Evidentemente, consolidaram-se em espaços midiáticos inexistentes aos que temos na atualidade, como os suportados pela internet e suas plataformas, as quais facilitam a consulta às ideias provenientes de intelectuais que norteavam o pensamento público e tinham lugar reservado nas páginas impressas (NASCIMENTO; ALVES, 2019).

Contudo, se podemos considerar como fato que as plataformas trazem aos historiadores da educação a possibilidade de consulta às fontes, também podemos afirmar que questões imperiosas se colocam, a saber: o tratamento da fonte, que durante muito tempo foi feito a partir da materialidade documental. Assim, há que se considerar que jornais e revistas devem ser pensados para além de repositórios de notícias. Para a História da Educação, a imprensa é vista como local singular da atuação pedagógica. Afinal, esse tipo de fonte é o espaço em que

intelectuais articulavam suas redes e lançavam projetos, a exemplo daquilo que, em inúmeros momentos da história da educação, julgou-se ser o mais adequado para reformas educativas.

Fosse nas páginas da imprensa educativa, nas folhas libertárias da imprensa anarquista, ou em revistas científicas, o que importa para a História da Educação não são apenas as informações sobre um novo método lançado ou sobre escolas inauguradas. Sendo assim, ao historiador da educação cabe a busca pela compreensão da hegemonia do pensamento educativo e da formação do consenso presente no discurso.

Antes mesmo da consulta à fonte, que na realidade atual pode ser localizada a um clique de distância, o historiador da educação deve ter em mente que sua tarefa está em captar os debates daqueles que se destacaram em discursos, contendas e projetos, os quais foram categorizados, inclusive, como intelectuais da educação.

Se projetos realizados no âmbito da pesquisa histórica podem sucumbir a tais percalços, o que dizer das investigações em educação realizadas em plataformas virtuais que atravessam espaços na simultaneidade do tempo? Ou mesmo da profusão das webconferências e das videochamadas que invadiram o espaço doméstico e escolar permitindo o deslocamento geográfico no encolhimento do tempo e do território, possibilitando “[...] estar em casa e simultaneamente no mundo externo com apenas um clique” (VIDAL, 2022: 5).

Não se deve ignorar que o cotejamento da fonte não pode ser feito apenas pela busca de palavras-chave, as quais podem inviabilizar os tensionamentos que buscaram consenso pela hegemonia discursiva. Por exemplo, um historiador da educação, calcado no estudo da imprensa e voltado a uma metodologia crítica, deve saber que a imprensa é produzida para a circulação de informações selecionadas e matérias escolhidas por aqueles que a detém, seja o editor, o grupo jornalístico que a mantém financeiramente, ou qualquer outro que tem o domínio daquilo que será publicado.

Ao considerar os elementos da produção do impresso, é possível observar que tais dados são importantes para o historiador da educação em seu trabalho investigativo. Até pouco tempo, a tarefa de pesquisar sobre a forma de impressão era algo presente na técnica metodológica inerente ao ofício. Ao analisar esse fenômeno, não se pode descartar o “rigor da materialidade”. Uma publicação impressa funciona em um tempo distinto se comparada a um post em uma plataforma digital. Questões que impactavam o jornal impresso eram caras ao pesquisador que utilizava a imprensa como fonte; tais elementos envolviam desde a produção material, como o

papel e a tipografia, até os anúncios financiados e a própria organização das matérias no periódico, como as manchetes, colunas ou notícias de página inteira. Nesse sentido, Cruz e Peixoto (2007) demarcam que a historicidade da imprensa não pode ignorar o percurso que vai desde a sua constituição até sua reinvenção. As autoras explicam que:

Questão central é a de enfrentar a reflexão sobre a historicidade da Imprensa, problematizando suas articulações ao movimento geral, mas também a cada uma das conjunturas específicas do longo processo de constituição, de construção, consolidação e reinvenção do poder burguês nas sociedades modernas, e das lutas por hegemonia nos muitos e diferentes momentos históricos do capitalismo (CRUZ; PEIXOTO, 2007: 257).

Ao desconsiderar à historicidade da imprensa, o pesquisador corre o risco de não adentrar em minúcias que antes eram essenciais, focando apenas no texto, sem a contextualização que outrora era fundamental. Nesses termos, o desafio metodológico carece, ainda, da seguinte indagação: como manter a densidade de uma análise quando o suporte deixa de ser o objeto físico e se transmuta para o texto digital? Este, por sua vez, é processado em plataformas que, muitas vezes, nem mesmo quem as utiliza compreende plenamente o funcionamento técnico subjacente. Todavia,

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 (BRASIL; NASCIMENTO, 2020: 201).

Postas as preocupações anteriores, é mister observar que o processo de transição de um suporte físico para o digital tende a fazer com que o historiador da educação se insira no campo das Humanidades Digitais. O termo vai além de sugerir apenas o uso dos mecanismos computacionais para acessar uma determinada fonte. Ele carrega consigo algo que se aproxima

de uma mudança epistemológica no tratamento dado às fontes, o qual ainda carece de mais estudos e é pouco aprofundado nos cursos de formação de historiadores.

Assim, traçando uma perspectiva de comparação, o que antes era pesquisado em acervos e arquivos para que depois, no trabalho interpretativo, fosse encontrado o “nó da trama”, na atualidade observa-se uma facilidade questionável, considerando que muitas vezes o referido “nó” é dado pelo modo como as interfaces e plataformas operam. O trabalho que anteriormente necessitava de um pesquisador lendo e relendo inúmeras páginas dentro das bibliotecas, hoje é modelado na esfera computacional. Até mesmo os discursos que outrora necessitavam de tempo para serem interpretados podem ser decodificados mais rapidamente no espaço do digital. Tão logo, as Humanidades Digitais estão para além da digitalização do acervo; elas se impõem, gerando novas camadas de interpretação arraigadas na mediação técnica das plataformas.

No momento presente, a diluição das barreiras da comunicação coloca-se como um dos desafios a quem busca analisar a imprensa, questão fundamental a ser discutida em um cenário no qual o periódico parece ter suas delimitações cada vez mais ampliadas. Exemplarmente, um periódico do século XIX que tivesse sua circulação restrita à Corte do Rio de Janeiro tinha como impacto direto a formação de ideias para aquele núcleo; assim, tanto o intelectual que escrevia para um jornal, quanto o próprio impresso em que publicou, possuíam destino certo. Nesse sentido, “Ao mesmo tempo que esse ambiente digital possibilita a ampliação de compartilhamentos de informações, ele também é capaz de restringir conforme queira, estando dependente das intencionalidades de quem a domina” (PEREIRA; DANTAS, 2025: 6).

A instantaneidade com que as ideias circulam nas redes digitais difere grandemente da circulação das ideias na imprensa física, principalmente no que se refere à celeridade, a qual impõe aos pesquisadores a busca da compreensão de como os bancos digitais funcionam, sobretudo quando se tornam uma espécie de “esfera pública digital”. Aliás, como categorizar, nesses termos, os buscadores que se tornam partícipes da pesquisa, considerando que localizam e selecionam periódicos e, ainda, separam o que é procurado pelas palavras-chave? Levanta-se, assim, outra preocupação: a neutralidade dos sistemas de metadados utilizados pelas plataformas. É sabido que as plataformas são construídas dentro da lógica de programadores e profissionais de outras áreas, como os da tecnologia da informação, sendo que a lógica utilizada pode tanto ser frutífera quanto prejudicial para a atuação do intelectual que o historiador investigará.

A partir do que foi exposto, evidencia-se que, em um universo de transformação, coloca-se outra função aos pesquisadores: a do rigor da materialidade, que antes era focado na especificidade da produção e circulação de uma imprensa física e agora se coloca na perspectiva de criar uma crítica da materialidade digital. Nesse âmbito, pode-se evocar com o uso das ferramentas digitais que estabelecem filtros e, muitas vezes, localizam apenas determinada parte de uma informação em um dado periódico, fazendo com que o pesquisador se afaste da visão de conjunto.

A densidade da análise, essencial para compreender a pesquisa histórica, sem se deixar levar-se por “*bit* e *bytes*, que concebe a realidade como reduzível a combinações e equações matemáticas” (VIDAL, 2022: 6). O desafio posto, então, é o de construir uma práxis que utilize a potência técnica das plataformas digitais sem abdicar do método histórico, garantindo que o algoritmo não apague a complexidade da pesquisa que se consolidou sobre impressos físicos. Obviamente, os novos formatos impostos pelas Humanidades Digitais alteram substancialmente a própria heurística histórica e a forma como os pesquisadores trabalhavam em seu ofício. Se, na pesquisa convencional, o historiador dependia da leitura folha por folha, o ambiente digital oferece a tentação do atalho e a imensidão de resultados.

O fato de uma plataforma retornar inúmeros resultados para uma busca sobre um projeto educativo não significa, necessariamente, que a complexidade daquele fenômeno foi capturada; porém, denota que o algoritmo localizou padrões de caracteres que coincidem com a consulta, deixando de fora as nuances, as entrelinhas e os silêncios que a máquina ainda é incapaz de ler. Sendo assim, é possível vislumbrar que a compreensão de como as ferramentas digitais auxiliam na interpretação da história é essencial para a formação do historiador.

Ferramentas distintas na engrenagem para a interpretação da História nas Humanidades Digitais

No contexto das Humanidades Digitais, novas terminologias se inserem, muitas delas distantes da linguagem praticada por historiadores, cientistas sociais, pedagogos e outros pesquisadores que desconhecem termos como OCR, utilizado para designar *Optical Character Recognition* (Reconhecimento Óptico de Caracteres), CAQDAS, sigla que significa

Computer-Assisted Qualitative Data Analysis Software, big data, software, entre outros. Essa terminologia faz parte de um novo letramento que se insere na lógica de um mundo digital ainda desconhecido para muitos. Pode-se entender que, “As Humanidades Digitais promovem o uso de ferramentas tecnológicas para a análise de fenômenos históricos e culturais, permitindo que os historiadores visualizem e interpretem dados de formas antes inimagináveis” (ALENCAR, 2024: 8).

Uma das marcas mais conhecidas para os historiadores a partir desse processo é a digitalização da imprensa periódica, fato que trouxe ganhos para a História da Educação. A possibilidade de acessar jornais publicados em diferentes locais sem a necessidade de deslocamento físico alterou profundamente o alcance das pesquisas comparativas. Contudo, essa expansão do acesso não elimina problemas; ela apenas os desloca. Como afirma Rollo:

Dito de outra forma, arquivos e bibliotecas verão seus conteúdos cada vez mais transformados e acrescentados em suportes digitais, crescentemente diversificados, devendo, para o efeito, criar infraestruturas, formar pessoas (não só as que lá trabalham, mas seus utilizadores), ampliando suas funções, constituindo repositórios digitais, sob pena de serem substituídos por outras entidades ou de perderem parte do patrimônio cultural dos contextos que pretendem representar. As instituições de ciência, tecnologia e ensino superior (para além dos arquivos e das bibliotecas que integram) terão de operar no mesmo sentido, cumprindo-lhes, porém, a missão de contribuir no plano científico e tecnológico e sobretudo no quadro da reflexão cultural e intelectual (ROLLO, 2020: 24).

Com o interesse de desvendar um pouco desse universo e das problemáticas dele decorrentes, dois termos serão discutidos com o intuito de demonstrar que os historiadores da educação talvez tenham de dominá-los, são eles: o OCR e o CAQDAS. O primeiro corresponde ao reconhecimento óptico de caracteres. Ele se estabelece como uma das tecnologias mais utilizadas para converter imagens em texto pesquisável. Em que medida essa tecnologia atua? A resposta simplificada demonstra que uma foto de um jornal, de uma revista ou de outro documento impresso pode ser transformado em texto pesquisável. Exemplarmente, quando localizamos um jornal presente na Hemeroteca Digital, ao abrir o documento conseguimos solicitar que ele busque determinado termo presente no jornal. Seguindo essa perspectiva, se

antes o historiador deveria ler todo o periódico para encontrar onde estava determinada informação, agora, ao utilizar determinada palavra-chave, ele pode ser direcionado diretamente ao local onde ela está situada no periódico.

Por sua vez, outra sigla que compõe a engenharia investigativa para auxiliar o ofício do historiador é a de CAQDAS. Ela remete aos softwares que, para além de realizar a busca por palavra-chave, também atuam fazendo interpretações, compilando dados, entre outros procedimentos. Brasil e Nascimento (2020), ao apresentarem uma análise sobre o termo, explicitam que:

Uma vez que o material que será analisado tenha sua origem em dispositivos digitais ou tenha sido convertido para o formato digital, é possível efetuarmos, dentro do ambiente digital desses programas, a totalidade das etapas da pesquisa qualitativa. Existem aplicativos com diversas funcionalidades e que são capazes de tratar dados nos mais variados formatos, seja em áudio, texto, vídeo ou imagem (BRASIL; NASCIMENTO, 2020: 214).

Os exemplos acima operam não só convertendo imagens em texto pesquisável ou apresentando dados informacionais; eles também fazem a seleção desses dados. É evidente que precisamos problematizar tais questões, até mesmo como parte da crítica pertinente ao historiador. Sabemos, por exemplo, que esses mecanismos dependem, entre outros fatores, da qualidade da digitalização, do estado de conservação do documento original e das tipografias utilizadas.

Em periódicos oitocentistas, ou em jornais impressos com recursos limitados, as falhas são frequentes. Palavras podem ser truncadas, nomes próprios podem ser mal reconhecidos e trechos inteiros tornam-se invisíveis aos mecanismos de busca. Nesse sentido, muitos dos buscadores podem utilizar formas variadas para realizar suas pesquisas, como, por exemplo, a procura por um periódico específico, em que o pesquisador teria acesso ao conteúdo do jornal com o objetivo posterior de fazer a análise; ou, então, a busca pelo conteúdo contido na plataforma por meio de termos isolados. Assim, por exemplo, se o historiador digita a palavra “Educação”, o sistema pode retornar assuntos variados, remetendo à esfera acadêmica que conceitua o termo ou ao uso coloquial da palavra em anúncios e notícias cotidianas.

De toda forma, o risco de se perder determinadas lógicas está presente, podendo, inclusive, ocorrer a perda de informações que constam em determinado jornal, mas que estariam diretamente vinculadas à pesquisa e que a máquina não consegue indexar. Exemplarmente, se pensarmos na grafia de termos pedagógicos usuais, como “método”, por vezes o algoritmo não reconhecerá variações recorrentes, como “metodo” ou “methodo”. Isso vale para “physica” – “física”, e outras grafias oitocentistas presentes em diversos documento. Sem dizer, naqueles que se apresentam mal-conservados, borrados e com manchas que são impossíveis de identificar a determinadas ferramentas óticas digitais. A inferência nos resultados das buscas, que apresentam ocorrências com falhas pode ser entendida na explicação abaixo:

[...] os resultados direcionam o acesso às ocorrências numa ordem específica – pautada pelo padrão do próprio site. Mesmo que o usuário classifique as ocorrências por título de acervos, muitas vezes teremos uma lista com sobreposição de períodos, visto que muitos periódicos têm acervos organizados em outros padrões diferentes daquele da estrutura de décadas que pauta a ferramenta de busca. Ao acessar um desses acervos para ler as ocorrências, o usuário será guiado pela estrutura e pela própria concepção técnica da ferramenta a seguir as ocorrências, mas não terá acesso organizado ao conjunto das ocorrências no interior do acervo e em relação ao todo dos resultados da busca (BRASIL, 2022: 193).

A argumentação apresentada acima demonstra que a facilidade da busca por termos isolados pode levar o pesquisador a ignorar aquilo que não está nas ocorrências. Nesses termos, os mecanismos de busca necessitam ser compreendidos em seu funcionamento. Por exemplo, quando o OCR falha ao processar uma palavra, seja por uma mancha do tempo ou por uma falha na digitalização, aquele trecho torna-se invisível para o mecanismo de busca. Cria-se, assim, o “silenciamento da fonte”, sem que se tenha maiores dados de onde o intelectual pode ter publicado um manifesto central para a sua rede de sociabilidade (BRASIL, 2022). Se o algoritmo não leu o seu nome corretamente, esse documento deixa de existir na prática, uma vez que toda a confiança do pesquisador foi depositada no buscador da plataforma utilizada.

Brasil (2022) auxilia na reflexão ao demonstrar que além dessa invisibilidade técnica, há a questão da fragmentação do discurso. A busca por palavras-chave isola o texto de sua

vizinhança editorial. Ao localizar um termo como “instrução pública” ou “modernidade pedagógica”, o sistema de busca entrega o recorte exato, mas oculta o que estava ao lado. Perde-se a noção de que aquele artigo de fundo dividia espaço com anúncios de remédios, notas de falecimento ou críticas teatrais, elementos que, no “rigor da materialidade” física, ajudavam a situar o lugar social do intelectual. Essa “leitura aos saltos” pode transformar o discurso intelectual em um dado abstrato, desprovido da ambiência política que Gramsci (1982) identificava como essencial para a formação da hegemonia.

Certo é que, a tecnologia não deve ser vista como um determinante isolado, mas como o desenvolvimento de processos que se inserem relações de pesquisas e interferem no método do pesquisador. Tais inferências estão presentes no âmbito epistemológico, mas não se pode negar que são próprias de mudanças que se dão nas próprias relações de produção. Não é possível pensar em inviabilizar que as mudanças tecnológicas auxiliem o historiador. Ao contrário, trata-se de benefícios que, pensados adequadamente, tornam suas funcionalidades benéficas e até necessárias ao trabalho investigativo. O algoritmo, OCR, e o CAQDAS, portanto, não são apenas ferramentas técnicas, mas componentes de uma nova infraestrutura que media a relação do historiador com a superestrutura de seus problemas investigativos.

O ofício do historiador no universo das Big Dates

As leituras que discutem o ofício do historiador no âmbito das Humanidades digitais apresentam que as ferramentas desenvolvidas pelas novas tecnologias tendem a demonstrar formas de pesquisa que agregam a atividade do historiador. Todavia, questionar os seus “usos” é algo salutar para que os “abusos” não se tornem um dificultador em universos em que o “fetichismo pelo dado” pode superar a informação ou o fato que circunscreve o trabalho do historiador (BRASIL, 2022).

As considerações que frisam a importância de pesquisadores conhecerem o universo das tecnologias que impactam em seu trabalho, apresentam que há um “letramento digital” a ser formado, permitindo que suas atividades ganhem mais corpo. Um outro termo que se impõe é o de “literacia algorítmica”. O termo indica que o historiador compreenda como as plataformas priorizam certos discursos em detrimento de outros, sob o risco de reproduzir uma visão de

mundo ditada pela arquitetura do *software*, e não pela realidade histórica. De acordo com Rosa, O conceito de literacia algorítmica tem sido “[...] estendido para designar as ‘habilidades e competências envolvendo a busca, a seleção, a análise, a avaliação e o processo da informação, considerando os meios, contextos e ambientes em que se encontra e se produz o conhecimento” (ROSA, 2016: 25).

Ser letrado, ou ainda, ter literacia algorítmica importa em saber que as ferramentas carregam interesses que ainda são obscuros. Assim, existe uma suposta neutralidade dessas tecnologias, fato que também se impõe como desafio metodológico. É preciso considerar que os sistemas de busca e a organização dos metadados nas Humanidades Digitais são construções humanas, atravessadas por escolhas políticas e técnicas. Estas escolhas estão presentes no leitor formado nesse universo. É prudente reafirmar que,

O leitor da era digital adquire novas formas de ler e, conseqüentemente, novas maneiras de apreender o conteúdo, diferente da leitura de um livro impresso, seja em um espaço físico mais reservado, fazendo com que ele esteja isolado numa leitura singular, ou em um ambiente aberto ao público, como no ônibus, no metrô, no shopping, em uma praça. Nesse caso, o leitor da era digital está voltado simplesmente à sua tela. Com isso, ele desenvolve novas habilidades e técnicas, ganhando suas características e peculiaridades, construindo a cultura do leitor de textos eletrônicos (NASCIMENTO; ALVES, 2019: 9).

As interfaces das hemerotecas digitais ditam, muitas vezes, os caminhos da pesquisa, privilegiando certos periódicos em detrimento de outros que, por questões técnicas ou financeiras, não receberam o mesmo tratamento de digitalização ou indexação. Logo, o historiador não está apenas diante de um arquivo; está diante de uma curadoria algorítmica que, se não for questionada, pode condicionar os resultados da investigação e limitar a compreensão das redes intelectuais. Portanto, o enfrentamento desse cenário exige que o pesquisador não se deixe seduzir pela eficiência do processamento. E aqui entra o ofício do historiador que, independentemente dos recursos utilizados para sua pesquisa, não pode ignorar as formas pelas quais suas fontes são alcançadas. Por isso, a mediação tecnológica não deve ser vista como uma simplificação do ofício, mas como uma nova camada de complexidade que exige vigilância epistemológica constante.

Se o *Big Data* oferece o panorama, cabe ao método histórico, e apenas a ele, oferecer a profundidade. O desafio não é apenas encontrar o documento que lhe servirá de fonte, mas compreender por que ele foi preservado, como foi digitalizado e de que maneira a plataforma está condicionando o encontro com esse fragmento do passado pedagógico. Essa “curadoria algorítmica” impõe um novo tipo de ceticismo ao pesquisador. Não se trata de negar o avanço que as Humanidades Digitais representam para a História da Educação, mas de compreender que a interface não é o espelho fiel do arquivo (LUCCHESI, 2014).

O que a plataforma entrega é um fragmento processado e o trabalho do pesquisador é que o transforma em dado e em história escrita e contada. Sem a consciência de que o algoritmo pode priorizar o que é tecnicamente mais “legível” em detrimento do que é historicamente mais “relevante”, o historiador corre o risco de construir narrativas enviesadas pela própria ferramenta. Nesse cenário de mediações técnicas, Anita Lucchesi (2014) defende a urgência de uma crítica das fontes digitais que dialogue com o rigor do ofício. Para a autora:

Uma vez individuados os aspectos que podem ser mais genericamente classificados como inéditos, caberia também interrogar como as novas dificuldades e oportunidades de pesquisa poderiam ser discutidas nos espaços de formação do profissional de história. Como discutir com graduandos e licenciandos essas questões? Cabem leituras, seminários, oficinas, experimentos? Se do lado da pesquisa se abrem todas essas ‘janelas’ com interrogações, perguntas se colocarão para o ensino? Ao mesmo tempo, em termos de divulgação científica, a inserção da história no ciberespaço também apresenta uma infinidade de ocasiões” (LUCCHESI, 2014: 51).

O uso de softwares como Sistemas de Informação Geográfica (SIG), por exemplo, não apenas facilita a análise de dados espaciais, mas também torna possível a criação de mapas históricos dinâmicos, que demonstram a relação entre tempo e espaço de maneira inovadora. Ao integrar essas novas tecnologias ao fazer histórico, os autores mostram como é possível conectar as investigações acadêmicas às questões práticas do ensino, produzindo uma História mais acessível e interativa (ALENCAR, 2024: 8).

A crítica à materialidade digital deve integrar o próprio procedimento metodológico. O documento digitalizado não é o equivalente transparente do impresso; ele é o resultado de um processo técnico que seleciona, organiza e, por vezes, silencia. Como afirma Barros, ao voltar-se

à imprensa, de modo especial aos jornais, é preciso ter em mente que os periódicos “são atravessados por posicionamentos em relação à realidade social, os quais se conectam visceralmente a certos interesses políticos, sociais e econômicos” (BARROS, 2022: 592).

O risco não reside apenas na falha técnica, mas na confiança excessiva depositada na ferramenta. Quando o historiador se apoia exclusivamente na busca por termos isolados, pode deixar de perceber a organização interna do periódico, suas hierarquias editoriais e a relação entre textos aparentemente desconexos. A leitura fragmentada tende a deslocar o discurso de sua ambiência original, empobrecendo a análise.

O que aqui se denomina “fetichização do dado”, partindo de leituras como as de Brasil e Nascimento (2020), Brasil (2022) e Vidal (2022), refere-se à tendência de atribuir ao resultado da busca um estatuto de evidência suficiente. No entanto, o dado digital é sempre mediado: resulta de escolhas técnicas, políticas de digitalização e critérios de indexação que nem sempre são transparentes ao usuário. O método histórico, nesse contexto, não se opõe à tecnologia, mas atua como instância de problematização.

Ademais, o ofício nesses novos tempos exige o enfrentamento de uma crença implícita de que a indexação total de periódicos e a aplicação de modelos de linguagem eliminariam a subjetividade e as lacunas da pesquisa histórica. Contudo, para o historiador da educação, a lacuna é também um dado histórico. Não se reconstroem os modelos formativos de um período apenas pelo número de periódicos educativos que foram digitalizados. Se o historiador se rende ao “fetichismo do dado”, aceita passivamente que o que não está digitalizado e indexado não existiu.

Em última análise, o desafio das Humanidades Digitais na História da Educação é ético e político. Algo que ainda necessita ser mais analisado por aqueles que se preocupam com o desenvolvimento da pesquisa histórica. É preciso sempre considerar a dubiedade que carrega a técnica da investigação historiográfica, ou seja, utilizar a escala do algoritmo para ampliar o horizonte da pesquisa, mas sem esquecer o olhar atento às entrelinhas que os algoritmos nos trazem.

Conclusão

A transição dos arquivos de papel para os algoritmos das Humanidades Digitais não representa o fim do método histórico, mas o seu maior teste de rigor na contemporaneidade. A digitalização da imprensa periódica e a mediação da Inteligência Artificial oferecem ao historiador da educação uma escala de investigação sem precedentes sobre os intelectuais e seus discursos. Contudo, essa ampliação de alcance traz o risco de análises superficiais, marcadas pelo imediatismo do dado e pela fragmentação da fonte. A técnica pode obscurecer a teoria se não houver vigilância epistemológica constante.

Nesse cenário, a formação do historiador da educação precisa contemplar aquilo que aqui se apresentou como literacia algorítmica. Não se trata de transformá-lo em cientista da computação, mas de sempre colocar em prática o ceticismo necessário para questionar as fontes, sobretudo, aquelas que se levantam pela engenharia das plataformas.

É preciso perguntar quais acervos foram silenciados para que outros fossem digitalizados e a quem serve a velocidade da busca quando ela negligência a densidade do contexto. Claro, questões próprias do desenvolvimento tecnológico são passíveis e ocorrem, mas, devem ser acompanhadas da crítica histórica. Para além, a pesquisa histórica a partir do universo digital não pode converter-se em novo mecanismo de exclusão, no qual apenas o que é tecnicamente processável seja considerado historicamente existente.

Diante disso, o estudo da imprensa na era digital exige o que podemos chamar de uma práxis historiográfica revista pela compreensão dos códigos presentes nas Humanidades Digitais. O pesquisador deve dominar a técnica sem se deixar dominar por ela, assumindo postura que ultrapasse a mera operação de sistemas de busca. As ferramentas de indexação e o processamento de grandes volumes de dados devem constituir ponto de partida, nunca de chegada. Mesmo quando o suporte assume a forma do algoritmo, a crítica a materialidade do periódico continua necessária.

O desafio para a História da Educação consiste em garantir que, entre papéis e algoritmos, a perspectiva analítica do sujeito, dotado de conhecimento para a investigação histórica, não seja apagada pela automação. A tecnologia deve ampliar as perguntas do historiador, não simplificar suas respostas.

A proposta de novas formas de trabalhos com história feita a partir do digital configura, portanto, um convite a repensar questões que ainda estão por ser estudadas, dentre elas, a ética do acesso à informação e as maneiras de se ter acesso a elas de forma ampla, sem os

silenciamentos que podem devir das ausências nas buscas por ocorrências. Somente uma articulação engajada entre o rigor do método histórico e as possibilidades das Humanidades Digitais permitirá preservar a densidade da pesquisa histórica, cara ao historiador da educação. É nessa tensão entre o arquivo físico e a interface digital que a História da Educação encontra sua renovação e mantém viva sua função crítica diante das novas mediações tecnológicas.

Referências bibliográficas

- ALENCAR, Breno Rodrigo de Oliveira (2024). Prefácio. In: COELHO, George Leonardo Seabra; PADOVAN, Regina Célia; RODRIGUES, Rodrigo Gouvêa (Orgs.). *Ensino e Pesquisa em História: Outros Temas e Abordagens*. Palmas, Tocantins: EDUFT, pp. 7-11.
- BARROS, José D'Assunção (2012). A fonte histórica e seu lugar de produção. *Cadernos de Pesquisa do CDHIS*. vol. 25, n. 2. Disponível em: <<https://seer.ufu.br/index.php/cdhis/article/view/15209>>. Acesso em: 12 jan. 2026.
- BARROS, José D'Assunção (2022). Considerações sobre a análise de jornais como fontes históricas, na sua perspectiva sincrônica e diacrônica. *História Unisinos*. São Leopoldo, vol. 26, n. 3, pp. 588-604, set./dez. Disponível em: <<https://revistas.unisinos.br/index.php/historia/article/view/21514>>. Acesso em: 12 jan. 2026.
- BRASIL, Eric (2022). pyHDB: ferramenta heurística para a Hemeroteca Digital Brasileira. *História da Historiografia*. Ouro Preto, vol. 15, n. 40, pp. 186-217, set./dez. Disponível em: <<https://www.scielo.br/j/hh/a/bSVsvXYTnksnHSyB4bDRrZb/?format=pdf&lang=pt>>. Acesso em: 19 nov. 2025.
- BRASIL, Eric; NASCIMENTO, Leonardo Fernandes (2020). História digital: reflexões a partir da Hemeroteca Digital Brasileira e do uso de CAQDAS na reelaboração da pesquisa histórica. *Estudos Históricos*. Rio de Janeiro, vol. 33, n. 69, pp. 196-219. Disponível em: <<https://www.scielo.br/j/eh/a/XNJJWhFFzPKdkhF6cyj5Bjv/?lang=pt>>. Acesso em: 10 jan. 2026.
- CRUZ, Heloisa de Faria; PEIXOTO, Maria do Rosário da Cunha (2007). Na oficina do historiador: conversas sobre história e imprensa. *Projeto História*. São Paulo, n. 35, pp. 253-270, dez.
- GRAMSCI, Antônio. *Os intelectuais e a organização da cultura*. 4 ed. Rio de Janeiro: Civilização Brasileira, 1982.
- LUCA, Tânia Regina (2008). A grande imprensa na primeira metade do século XX. In: MARTINS, Ana Luiza; LUCA, Tânia Regina de (Orgs.). *História da imprensa no Brasil*. São Paulo: Contexto, pp. 149-175.
- LUCCHESI, Anita (2014). Por um debate sobre História e Historiografia Digital. *Boletim Historiar*, n. 2. Disponível em: <<https://seer.ufs.br/index.php/historiar/article/view/2127>>. Acesso em: 10 jun. 2022.
- NASCIMENTO, Ester Fraga Vilas-Bôas Carvalho do; ALVES, Josué dos Santos (2019). História da educação, humanidades digitais e divulgação científica: a biblioteca digital de História da

- Educação. *Cadernos do Tempo Presente*. São Cristóvão, vol. 10, n. 1, pp. 12-24, jan./jul. Disponível em: <<http://www.seer.ufs.br/index.php/tempo>>. Acesso em: 22 fev. 2026.
- PEREIRA, Débora Cristina de Souza; DANTAS, Lidemberg Régis Santos (2025). O uso das fontes na História Digital da Educação. *Revista Conexão ComCiência*. vol. 5, n. 1. Disponível em: <<https://revistas.uece.br/index.php/conexaocomciencia/article/view/14476>>. Acesso em: 12 jan. 2026.
- RIBEIRO, Ana Paula Goulart (2015). Nelson Werneck Sodré e a história da imprensa no Brasil. *Intercom – Revista Brasileira de Ciências da Comunicação*. São Paulo, vol. 38, n. 2, pp. 275-288, jul./dez. Disponível em: <<https://www.scielo.br/j/intercom/a/LDp7RKsgqbpP6Kd6zxPZCtq/?format=pdf&lang=pt>>. Acesso em: 26 nov. 2025.
- ROSA, Beatrice Bonami (2016). *A transdisciplinariedade das literacias emergentes no contemporâneo conectado*: um mapeamento do universo documental das literacias de mídia e informação (MIL). Dissertação (Mestrado em Interfaces Sociais da Comunicação). Escola de Comunicações e Artes, Universidade de São Paulo, São Paulo.
- ROLLO, Maria Fernanda (2020). Desafios e responsabilidades das humanidades digitais: preservar a memória, valorizar o patrimônio, promover e disseminar o conhecimento. O programa Memória para Todos. *Estudos Históricos*. Rio de Janeiro, vol. 33, n. 69, pp. 19-44, jan./abr. Disponível em: <<https://doi.org/10.1590/S2178-149420200001000003>>. Acesso em: 13 fev. 2026.
- SODRÉ, Nelson Werneck (1999). *História da imprensa no Brasil*. 4. ed. Rio de Janeiro: Mauad.
- VIDAL, Diana Gonçalves (2022). Humanidades digitais e cultura material (escolar). *History of Education in Latin America – HistELA*. vol. 5. Disponível em: <<https://periodicos.ufrn.br/histela/article/view/30136>>. Acesso em: 25 nov. 2025.

Declaração de Uso de Inteligência Artificial Generativa

Os autores do artigo acima identificado declaram que ao longo da elaboração deste manuscrito foi utilizada a seguinte ferramenta de inteligência artificial: Gemini, com a finalidade de: revisão de linguagem — aprimoramento do estilo textual e da clareza; apoio ao levantamento preliminar de referências bibliográficas sobre a temática do estudo; organização inicial de literatura; identificação de autores, conceitos e obras potencialmente relevantes; auxílio na estruturação de tópicos e na sistematização de informações. Após o uso das ferramentas, os autores revisaram e editaram criticamente todo o conteúdo, assumindo total e integral responsabilidade pelo manuscrito publicado, inclusive por eventuais imprecisões ou incorreções.