



A AUTENTICIDADE DO INAUTÊNTICO: ARQUIVOS DIGITAIS NO ESTUDO DO TEMPO PRESENTE

The authenticity of the inauthentic: digital files in the study of the presente

La autenticidad de lo no auténtico: los archivos digitales en el estudio de la actualidad

André Ricardo Valle Vasco Pereira¹

Resumo: Este trabalho lida com o problema da autenticidade de arquivos digitais aplicados ao estudo de História do Tempo Presente. Ele busca distinguir arquivos com elementos que foram digitalizados de arquivos que foram criados originalmente como digitais. Distingue arquivos digitais individuais de séries documentais digitais. Defende que o conceito de autenticidade dos arquivos digitais deve ser concebido em um continuum, de um maior a menor grau. Diferencia repositórios fonte e destino dos arquivos digitais. Discute o fato de que muitos acervos de arquivos digitais não foram feitos para serem oferecidos por tempo indefinido. Aborda o caráter compósito dos arquivos digitais, com elementos temporários, o que dificulta o seu *download* completo. Leva em consideração que os custos de armazenamento e de heurística deste tipo de documento não devem ser desprezados. E avalia os problemas de variabilidade dos arquivos digitais.

Palavras-chave: Autenticidade. Arquivos Digitais. Séries Documentais Digitais.

Abstract: This article deals with the problem of authenticity of digital files applied to the study of History of the Present. It seeks to distinguish files with elements that have been digitized from files that were originally created as digital ones. Distinguishes individual digital files from digital source series. It argues that the concept of authenticity of digital files must be conceived on a continuum, from a greater to a lesser degree. Differentiates source and destination repositories of digital files. Discusses the fact that many digital file collections were not designed to be offered indefinitely. It addresses the composite nature of digital files, with temporary elements, which makes their complete download difficult. Take into account that the storage and heuristic costs of this type of document should not be ignored. And it evaluates the variability problems of digital files.

Keywords: Authenticity. Digital Files. Digital Source Series.

Resumen: Este trabajo aborda el problema de la autenticidad de los archivos digitales aplicado al estudio de la Historia del Tiempo Presente. Busca distinguir archivos con elementos que han sido digitalizados de archivos que fueron creados originalmente como

¹Doutor em Ciência Política pelo IUPERJ. Professor do Departamento de História da Universidade Federal do Espírito Santo, Vitória, Espírito Santo, Brasil. E-mail: andre.r.pereira@ufes.br; ID Lattes: <http://lattes.cnpq.br/2814186365798017>; ORCID iD: <https://orcid.org/0000-0002-9400-7611>.

digitales. Distingue archivos digitales individuales de series documentales digitales. Sostiene que el concepto de autenticidad de los archivos digitales debe concebirse en un continuo, de mayor a menor grado. Diferencia los repositorios de origen y destino de archivos digitales. Analiza el hecho de que muchas colecciones de archivos digitales no fueron diseñadas para ofrecerse indefinidamente. Aborda el carácter compuesto de los archivos digitales, con elementos temporales, lo que dificulta su descarga completa. Tenga en cuenta que no se deben ignorar los costos de almacenamiento y heurísticos de este tipo de documentos. Y evalúa los problemas de variabilidad de los archivos digitales.

Palabras clave: Autenticidad. Archivos Digitales. Serie Documental Digital.

Introdução

Este trabalho trata do problema da autenticidade de arquivos digitais na pesquisa em História do Tempo Presente. Para tanto, lança mão de um conjunto de experiências de seu autor, adquiridas ao longo de sua trajetória profissional. De forma mais específica, o que se verifica é que, muitas vezes, os arquivos digitais mais utilizados neste tipo de pesquisa apresentam questões de autenticidade, que precisam ser endereçadas.

Logo de início, informo que não tenho o propósito de oferecer uma definição indiscutível de tempo presente, o que, em si, é problemático (BRUNI; MENNA-BARRETO; MARQUES, 2007). Basta dizer que arquivos digitais são parte importante da pesquisa em temas com recortes temporais mais recentes, sendo este o caso que desejo considerar. Quando tomamos concepções rigorosas de autenticidade para documentos deste tipo, a conclusão, muitas vezes, é a de que estamos utilizando fontes “inautênticas”.

Ora, na pesquisa em História, é comum que tenhamos que lidar com problemas empíricos e produzir explicações, mesmo que a base documental apresente limitações insanáveis. Neste caso, o que eu pretendo sugerir é que **os acervos contendo arquivos digitais individuais ou séries documentais digitais podem apresentar graus variáveis de autenticidade**. O ideal é o de buscar lidar com o máximo de autenticidade possível, o que passa vários tipos de problemas, conforme pretendo desenvolver daqui para frente.

Discutindo autenticidade

Jonathan Domingues atenta para a compreensão de como são organizados os acervos de arquivos digitais (diretórios e inventários). Citando a obra de Duranti (1994), afirma, sobre o conceito de autenticidade, que: “os arquivos são criados como verossímeis e confiáveis para

quem os necessita para agora. São mantidos com garantias apropriadas para ação futura e para informação” (DOMINGUES, 2023, p. 20).

Ainda são consideradas duas outras características: naturalidade e organicidade, as duas se referindo a processos de coleta, acúmulo progressivo e coeso dos arquivos. O autor afirma que não há definição fechada sobre o termo autenticidade, mas observa que: “A autenticidade dos documentos digitais só pode ser sustentada se eles estiverem seguros” (DOMINGUES, 2023, p. 21). Adiante, citando Rondinelli (2005), surge o seguinte:

o conceito de autenticidade refere-se à adoção de métodos que garantam que o documento não foi adulterado após a sua criação e que, portanto, continua sendo tão fidedigno quanto era no momento em que foi criado (DOMINGUES, 2023, p. 21).

O autor lida com um repositório institucional. E observa que o *software* utilizado não contempla os recursos obrigatórios de manutenção da autenticidade, de forma que um repositório não é automaticamente confiável deste ponto de vista. Nesta discussão, um aspecto a mais citado é que a autenticidade deve ser garantida a longo prazo. Seguindo Gilliland-Swetland, diz que pode haver “comprometimento da autenticidade do arquivo em sua completude, em virtude da durabilidade dos armazenadores, dos *softwares*, *hardwares* e dos próprios arquivos digitais” (DOMINGUES, 2023, p. 28)

Já na parte empírica do trabalho, o autor observa que, com o tipo de arquivo que lidou, a nota de rodapé (com o *link* do arquivo usado) foi o meio que considerou adequado para garantir a autenticidade. Outra possibilidade é dar orientações a quem lê sobre como encontrar as informações, no texto, em notas ou em apêndices. As questões levantadas são relevantes. Eu gostaria, contudo, de destacar outros aspectos que surgiram da minha experiência.

Ingresso e permanência dos dados

O que emerge da discussão colocada por Domingues e interessa à reflexão que está sendo feita aqui é: a organização dos acervos digitais é fundamental para discutir sua autenticidade. Um dos aspectos é o ingresso dos materiais. Outra é a sua permanência e acesso ao longo de um prazo indefinido. Neste sentido, cabe discutir os componentes do acervo. No caso do trabalho de Domingues, foram textos. No meu caso, são:

- a) matérias jornalísticas publicadas por mídia *online* corporativa (e outras, que as imitam ou simulam);
- b) matérias jornalísticas publicadas por organizações que possuem meios próprios de comunicação (como jornais de sindicatos de trabalhadores ou patronais);
- c) postagens feitas em redes sociais;
- d) documentos de tipos variados (textos, planilhas, dados organizados por meio de *softwares* pouco comuns), publicados por entidades públicas ou de direito privado.

Quando pensamos nestes tipos de fontes primárias, um aspecto que se destaca é o seu **caráter** muitas vezes **compósito**. É possível que uma matéria de jornal *online* possua apenas texto como seu conteúdo. Todavia, em geral, não se trata de um elemento apenas. Um mesmo arquivo digital pode conter: texto, imagens de variados tipos, áudio, vídeo, *links*, documentos anexados. As imagens por sua vez, podem ser:

- a) fotografias;
- b) gráficos;
- c) tabelas que foram inseridas como anexos e não como parte do texto (que são navegadas por meios próprios);
- d) recursos bem variados de ilustração (ícones; letras especiais destacadas; desenhos; grafismos; imagens dinâmicas, como *GIFs*; etc).

Os áudios inseridos podem fazer parte dos vídeos ou não. Podem ser trechos de gravações, como no caso de conversas de celulares obtidas com permissão legal. Podem ser músicas. Podem ser efeitos ou trilhas sonoras. Neste caso, se o arquivo é acessado com o som do dispositivo (computador, celular) desligado, nem chegam a ser notados no momento da coleta. Vale notar que muito do que foi apontado é exibido apenas no momento do acesso. De forma mais específica, os sites costumam publicar propagandas, que são programadas para atender a critérios que algoritmos capturam dos IPs (endereço do protocolo de *internet*) de quem está navegando. Ou seja, há exibições de *links*, por exemplo, que variam de acordo com quem busca a matéria *online*.

Quando uma fonte primária deste tipo é selecionada para sofrer *download* e constituir um elemento em um acervo digital, o mais correto seria que **toda a composição** fosse salva e preservada da mesma forma que foi publicada originalmente. Assim como, ao ser acessada posteriormente no acervo, deveria ser reconstituída da mesma maneira que foi exibida. Há aí dois problemas. O primeiro é o fato de que baixar toda a composição é mais caro que apenas

partes selecionadas dela. É mais caro porque leva mais tempo e exige mais espaço de armazenamento. O segundo problema deriva da dificuldade em salvar todos os elementos citados. *Links* dinâmicos direcionados por algoritmos não podem ser guardados. E há muitos componentes da fonte original que foram projetados por seus *webdesigners* para que não pudessem ser salvos *offline*.

Salvo erro de julgamento da minha parte, não há repositórios, públicos ou privados, direcionados à pesquisa em História do Tempo Presente que ofereçam acervos de jornais *online* com as características citadas acima, ou seja, que salvem TODOS os conteúdos e os ofereçam, de forma pública e gratuita, com a garantia de acesso por tempo indeterminado. Um problema bem simples, por exemplo, está na preservação do URL (*Uniform Resource Locator*), quase sempre reconhecido por iniciar com *www* (mas nem sempre) que direcionava à matéria. Muitos foram desativados. Mais grave, entretanto, é a simples ausência desta informação, como no caso, digamos, do aplicativo *Internet Archive* (<https://archive.org/>).

Desta forma, **um dos critérios de autenticidade** apresentados por Domingues, **que é justamente o link, nem sempre pode ser apresentado** ou, quando foi salvo na coleta original da fonte, pode estar desativado. Neste caso, o autor sugere a indicação do caminho tomado para a obtenção do material, o que, na prática recai na indicação do URL ou nome geral do site/plataforma.

Falando de jornais *online*, em alguns casos, eles oferecem acesso a seus acervos, inclusive com edições já digitalizadas. Só que isto não é uma regra. Não há nenhuma garantia de que estes periódicos mantenham, por tempo indeterminado, seus acervos disponíveis. E há publicações que cobram pelo acesso de todo ou parte do material. Neste sentido, a monetarização completa ou parcial de acervos nos leva a introduzir **outro critério para a autenticidade que é a gratuidade**. Na medida em que ela não seja garantida, o acesso fica comprometido.

Além disso, a monetarização das plataformas nas quais buscamos fontes primárias também passa pela captura de informações de navegação de quem faz a pesquisa, sendo, muitas vezes, obrigatório concordar com termos de uso que legalizam tal acesso. Esses dados são usados para a exibição de propaganda direcionada. E surge daí outra questão que é a venda de tais dados. Nos acervos físicos, quem pesquisa se identifica e os gestores das instituições costumam usar tais informações para prestação de contas.

Todavia, é diferente a política de monetarização das plataformas, pois, invés de só oferecer bases de dados, elas podem usar estes acervos para capturar informações de navegação dos usuários das quais possam lucrar. Existe, neste sentido, a potencialidade de que materiais sejam selecionados pelas empresas que são deles proprietários para serem **disponibilizados** justamente **por oferecerem maior retorno econômico**. Esta seleção é muito grave, pois a autenticidade depende da exibição de TODO o acervo. Esta é a única forma de evitar vieses. Se o detentor dos documentos faz qualquer tipo de seleção, a autenticidade fica comprometida.

Vale notar que certas priorizações não chegam a ser incomuns, mesmo em instituições públicas. O Arquivo Público do Estado do Espírito Santo (APEES), por exemplo, destaca listas de imigrantes que ingressaram em terras capixabas. No momento em que este texto está sendo redigido, o seu site faz isso de forma mais discreta que no passado (<https://ape.es.gov.br/>). A razão para o destaque está no fato de que muitos descendentes de europeus usam os documentos para obter cidadania, em especial na Itália. Há mesmo empresas que prestam tal serviço e são usuárias do serviço. Neste caso, não há problemas de autenticidade, mas é uma ilustração de exibição pública de parte do acervo, de um grau bem moderado de viés na apresentação. Se isto ocorre com um repositório confiável como este, o que falar das empresas? Existem jornais que, na cobertura diária, exigem identificação e/ou assinatura ou bloqueiam certos conteúdos para quem não o faz. Mais tarde, em seus arquivos “abertos”, mantém os mesmos critérios.

Acima, foi observado que matérias de jornais *online* (assim como muitas publicações de internet que também são possíveis fontes primárias) são compósitos. Ao serem salvos e exibidos posteriormente, muita coisa pode ser perdida, sendo particularmente relevantes as imagens, gráficos e tabelas que complementam ou interagem com o texto, que é mais fácil de ser salvo e exibido depois. Este ponto será mais desenvolvido adiante. E para que a autenticidade seja garantida, mesmo em repositórios confiáveis, não há segurança de recuperação de todos os componentes originais.

Adicionalmente, há outra questão correlata, que remete a processos de digitalização dos documentos, presente também em repositórios confiáveis. O ponto aqui é: ao ser inserido na internet um elemento empírico, feito originalmente para o meio virtual ou não (textos em papel, quadros, afrescos, estátuas, documentos de cultura material etc.), a autenticidade depende da fidelidade ao original na exibição. Digitalizações malfeitas, com equipamentos de

baixa resolução, certas vezes sem o recurso de reconhecimento ótico de caracteres, OCR (no formato de imagem), entregam textos ilegíveis e imagens imperfeitas. Não é incomum, por exemplo, na Hemeroteca Digital da Biblioteca Nacional (<https://bndigital.bn.gov.br/hemeroteca-digital/>), que periódicos sejam apresentados em pdfs com estes problemas.

Ademais, fotografias de quadros, digamos, costumam ser exibidas em baixa resolução e quase sempre, só de frente. Para quem faz pesquisa com iconografia, não é impossível que haja interesse em saber a profundidade do quadro, marcações de suportes, possíveis elementos na parte traseira. Este ponto é mais grave para esculturas, que não costumam ser exibidas em fotos de 360° ou em diferentes ângulos. Outra situação: quando falamos de inscrições em uma coluna, fotos na sua altura são sempre muito válidas e necessárias, mas também é importante, se possível (caso a coluna permaneça em pé), que sejam oferecidas fotos da perspectiva de quem se encontra na base. Por quê? Porque esta foi a visão que o público teve das inscrições na época em que foram criadas. Se for aceito que arquivos digitais de imagens só podem ser autênticos caso atendam a tais critérios, praticamente nada se encaixa. Este problema é tão intrincado quando pensamos na realidade dos repositórios que o conceito de autenticidade precisa incorporar uma noção de grau. Ou seja, uma foto em mais baixa resolução tem menor autenticidade que outra, com maior resolução. E o ideal é que os repositórios busquem alcançar maior autenticidade, de acordo com seus meios (sempre escassos).

O custo dos acervos e a autenticidade

O ponto que acabou de ser indicado acima remete ao problema dos custos. A montagem e manutenção dos acervos digitais implica várias despesas, entre elas a remuneração da mão de obra que faz a tarefa. Quanto menos especializada e mais mal paga, menor interesse e compromisso de quem trabalha em lidar com os problemas citados. E há custos dos provedores de rede e de armazenamento. Neste último quesito, uma dificuldade se soma: a mudança constante dos padrões industriais dos dispositivos, com o surgimento de novas mídias e o abandono das anteriores (COSTA; PINTO, 2017). Sobre este ponto, farei um relato pessoal para servir como ilustração das dificuldades. Adquiri meu primeiro microcomputador em 1991, graças a uma bolsa da Fundação Ford, com o objetivo de redigir

minha dissertação de mestrado em História. Tratava-se de um PC XT, rodando o sistema operacional MS-DOS (*Microsoft-Disk Operating System*).

Aqui se abre uma janela para um comentário sobre os sistemas operacionais. Na medida em que eles mudam ao longo do tempo e competem entre si, geram sérios problemas de compatibilidade entre arquivos e, com isso, dificuldade ou, até, na prática, a impossibilidade de que certos arquivos digitais sejam recuperados e exibidos em acervos, mesmo que tenham sido adequadamente salvos na época. O MS-DOS não era o único que servia a microcomputadores que fossem clones de IBM PC, como era o caso dominante no mercado brasileiro nos anos 1980 e início dos 1990. Arquivos que tenham sido produzidos no ambiente DR-DOS, por exemplo, não podiam ser lidos na interface MS-DOS e vice-versa. Além disso, microcomputadores da *Apple* sempre usaram soluções proprietário de sistema operacional.

Com o surgimento da interface *Windows*, em 1985, mas com efetiva popularização a partir dos anos 1990, passou a existir um esforço para que arquivos produzidos por diferentes *softwares* e sistemas operacionais pudessem ser convertidos e/ou lidos. Este foi, contudo, um processo. Durante muito tempo, dados gerados em diferentes sistemas operacionais só rodavam neles. E os sistemas foram se tornando obsoletos, descontinuados ou derrotados pela competição. Isto significa que **existe hoje uma massa de arquivos digitais**, que são de interesse para a pesquisa em História, e **que não pode ser acessada na prática** ou exige *expertise* profissional muito cara.

Os arquivos citados acima foram preservados em meios de arquivamento que, por sua vez, passaram a ser abandonados. No caso do meu exemplo, o microcomputador era equipado com *drives* de disquetes de 5 ¼ polegadas. Além disso, possuía um disco rígido (HD) interno. Ambos eram formatados segundo o padrão MS-DOS. Hoje, não tenho condições práticas e acessíveis para recuperar os arquivos que produzi na época.

Lembrando que, por se tratar de um momento anterior à internet (ou seu modelo mais disseminado, o *www*), os arquivos digitais de programas e de dados pré-gravados eram disponibilizados apenas por meio de disquetes. E os arquivos gerados nos microcomputadores eram salvos nos disquetes e nos HDs. O disquete de 5 ¼ foi criado em 1976. Em meados da década de 1980, começaram a ser substituídos pelos disquetes de 3 ½ polegadas. Estes, contudo, teriam que ser lidos por *drives* específicos. Ou seja, os microcomputadores que tinham *drives* mais antigos ficaram reféns do padrão anterior, que foi sendo eliminado.

No caso do Brasil, devido ao atraso tecnológico e do alto custo dos equipamentos, que eram cotados em dólar, numa fase de hiperinflação, mesmo quem conseguia comprar um computador tendia a adquirir máquinas já ultrapassadas. As empresas estatais e entidades públicas brasileiras da época eram proibidas por lei de comprar equipamento mais moderno, pois tinham que dar preferência ao “equivalente” nacional. Com isso, muitos arquivos digitais privados ou públicos criados na época não estão em boas condições para serem inseridos em repertórios. Ou exigem um custo alto e impraticável para serem recuperados, adaptados e disponibilizados.

Lembrando que alguns dos materiais são privados, como o meu, e, por diversas razões, podem ser de interesse para preservação. E houve empresas ou entidades públicas que continuaram operando com equipamentos e programas obsoletos por muito tempo, entre outras razões, para não gastar com treinamento de pessoal e alterar rotinas de trabalho. **Toda uma produção de arquivos digitais** que deve ter se iniciado nos anos 1960, em computadores *mainframe*, até a popularização na internet e da interface *Windows*, **se encontra fora do alcance dos acervos.**

As mudanças nos padrões de arquivamento estiveram acompanhadas da introdução de novos dispositivos de entrada e saída. Os disquetes de 5 ¼ tinham *drives* específicos, assim como os de 3 ½, como já assinalado. Em 1982, foi criado o *Compact Disc*. Em 1984, foram introduzidas especificações para que ele pudesse gravar dados digitais. A seguir, passaram a ser vendidos *drives* de leituras de CDs que podiam ser apenas lidos (CD-ROM). Com o tempo, surgiram outros tipos de CDs, graváveis e *drives* que também podiam gravar. Após o ano 2000, começou o declínio do uso de CDs como mídia de arquivamento, até que os computadores novos pararam de ser vendidos com *drives* para eles.

Em 1994, foi criado o *Zip Drive*, que surgiu como uma opção que parecia ser uma boa alternativa aos disquetes, ainda mais se o usuário não tivesse um leitor de CD. As mídias de *Zip Drive* ofereciam mais espaço que os disquetes. Quem apostou nelas, como eu, se viu prejudicado, pois a empresa proprietária da mídia foi processada por defeitos de fabricação e descontinuou o produto. Fiquei com uma grande quantidade de *zip disks* que não consigo ler. A solução foi migrar para os CDs, que, mais tarde, também foram descontinuados.

Na atualidade, discos rígidos externos e *pen drives* de alta capacidade são a opção, mas a narrativa que vem sendo apresentada nos levará a concluir que eles também serão

substituídos. A opção por eles foi forçada pela indústria, entre outras razões, pelo fim dos *drives* de CD e a imposição dos conectores USB como residentes nas máquinas. Atualmente, o armazenamento em nuvem vem sendo apresentado como a nova fronteira. Precisamos ter em mente que, ao longo do período que venho considerando (dos anos 1980 até 2020), arquivos de áudio, de vídeos e imagens compostas de tipos variados cresceram muito de importância no universo dos materiais digitais. Eles exigem muito espaço de armazenamento. Em geral, não podem ser anexados a e-mails. Este ponto em particular obriga os usuários a buscarem plataformas com arquivamento em nuvem, o que implica custos significativos e dependência de políticas estabelecidas pelas empresas que prestam este serviço.

Durante a pandemia de Covid-19, por exemplo, a Universidade Federal do Espírito Santo fez um contrato para adquirir serviços do *G Suite for Education*, que fornecia ambiente virtual para aulas e espaço para salvar transmissões gravadas ou ao vivo. Tais arquivos eram pesados e foram se acumulando, para que ficassem a disposições dos discentes. Em determinado momento, contudo, a empresa *Google* determinou um limite muito baixo de manutenção dos arquivos, que tiveram que ser apagados. Este exemplo nos mostra que dificuldades de armazenamento por usuários privados e públicos nem sempre são solucionados pelo ambiente nuvem. Erros e esquecimento de prover o *backup* adequado também podem ser acompanhados por defeitos que os equipamentos sofrem. Já passei pela situação de ter HD externo perdido. Todos os problemas citados contribuem para que um volume significativo de arquivos digitais: a) não sejam salvos; b) sejam salvos, mas não possam ser recuperados; c) estejam disponíveis em alguma mídia, mas o custo de sua reprodução em algum repertório seja proibitivo; d) que alguns repertórios sejam disponibilizados, mas sem a devida continuidade, o que compromete de forma direta a autenticidade, conforme discutido por Domingues.

Repositórios fonte e destino 1

Há reflexões interessantes sobre repertórios de arquivos digitais criados a partir de documentos em papel que foram digitalizados (SOUZA, 2023) ou objetos e imagens que foram fotografados. O artigo de Arthur Ferreira Reis discute o campo das Humanidades Digitais, focando em programas que elaboram análises quantitativas do discurso. Ele observa que: “iniciativas governamentais de digitalização e disponibilização on-line de documentos [...] colaboraram para que pesquisadores construíssem seus próprios repositórios” (REIS,

2023, p. 38). A partir daí, cita alguns exemplos. O que temos aqui? Fontes iniciais de dados digitais/digitalizados que alimentam a formação de repertórios por parte de pesquisadores.

No caso do estudo do autor, são citados acervos que acabaram sendo disponibilizados ao público (o que é diferente de um acervo pessoal privado feito para a pesquisa, como é o meu caso). Entre os casos abordados, ele cita uma produção de sua lavra, reunida no site <https://jornaisdaindependencia.com.br/>. Trata-se de um recurso que foi financiado pela Fundação de Amparo à Pesquisa do Espírito Santo (FAPES) em conexão com um laboratório do programa de pós-graduação em História da UFES. Ali, foram reunidos jornais que circularam durante a independência do Brasil e que podem ser baixados. O referido material foi coletado de diferentes fontes e hospedado no site.

O projeto também inclui uma plataforma de análise textual, assunto que é discutido pelo autor no artigo. No momento do acesso (27/09/2023), contudo, o *link* para este recurso não parecia estar funcionando (<https://jornaisdaindependencia.com.br/analises-textuais/>). De qualquer forma, o que o projeto consolida é um tipo de disseminação de arquivos digitais que está ocupando algum espaço. Ele existe a partir de uma oferta original de arquivos por parte de outros meios, com o detalhe de ser orientado a um propósito específico, configurando um campo determinado de estudo. Ou seja, há **repositórios fonte** e **repositórios destino**, que reordenam o material. Neste caso e no próximo que será citado, o trabalho é financiado com recursos públicos.

Há repositórios de arquivos digitais pertencentes a instituições públicas, a instituições sem fins lucrativos (em geral com recurso público) e a empresas que os utilizam, conforme abordado acima, com o fito de obter lucros. Eles são diferentes entre si. Não é impossível que haja viés na oferta *online*, como foi comentado quanto ao tema nos imigrantes no APEES. No caso de empresas, o viés é definido pelo lucro, o que é mais grave. Pois bem, os três tipos citados de repositórios fonte é que alimentam os repositórios destinos, nos quais, a partir de critérios definidos por seus criadores, os arquivos digitais são reorganizados. Pode haver, portanto, viés na fonte e no destino, colocando em questão não a autenticidade do arquivo digital individual, mas do acervo, no sentido de se garantir a oferta de tudo o que se dispõe.

O interessante exemplo citado acima merece ser comparado com outro do mesmo tipo (repositório destino), que entrou no ar em 26 de setembro de 2023 (SCHUINA, 2023), intitulado “Portal Patrimônio Capixaba” (<https://www.patrimoniocapixaba.com/>). Ele contou

com financiamento do Fundo de Cultura do Estado do Espírito Santo (Funcultura). Foi elaborado por uma equipe liderada por uma professora de história com o objetivo de fornecer materiais para o ensino e para o público em geral, com foco no patrimônio material e imaterial capixaba. O que se fez foi reunir certa quantidade de textos, de imagens e de alguns recursos didáticos a partir de repertórios fonte, disponibilizados na nuvem para download.

Esse exemplo busca um público mais amplo que o anterior. Qual é a questão? Uma proposta de origem privada, que disputou um edital público, foi contemplada, gerou um repositório destino, e usa o arquivamento em nuvem. Uma das características da autenticidade é sua oferta por tempo indeterminado, o que poderá ser prejudicado caso a empresa que hospeda os materiais em nuvem deixe de ser paga, o que é uma possibilidade, conforme já observado antes. Por fim, vale notar que estas bases de dados, é claro, são de grande importância, porém se trata de um tipo de iniciativa que lida com problemas diferentes de reunião e disponibilização de arquivos originalmente digitais e de recuperação muito difícil. Ou seja, tem sido mais comum formar repertórios com tais fontes, deixando uma lacuna no outro tipo.

Repositórios fonte e destino 2

Ao participar da equipe que redigiu um livro sobre as intervenções nas universidades federais durante o governo Bolsonaro (PEREIRA; ZAIDAN; GALVÃO, 2022), percebi ser necessário recolher arquivos digitais. Foram usados documentos *online* de vários tipos: jornais, portais de notícias, *blogs*, sites regionais de opinião, muitas vezes elaborados sem critérios confiáveis de profissionalismo.

A estratégia seguida foi necessária para obter informações sobre detalhes acerca da escolha dos reitores considerados “interventores”. Os atores envolvidos, por razões evidentes, não tinham interesse em se apresentarem com formas de identificação que pudessem afetar sua legitimidade. A busca por um grau mínimo de objetividade e a impossibilidade de realizar levantamentos presenciais em todas as instituições conduziu ao critério de buscar pelas informações disponíveis na internet, mesmo que pudessem ser oriundas de bases de dados que são, muitas vezes, tendenciosas. Isto gerou questões relevantes para o trabalho com arquivos digitais na História do Tempo Presente em relação a temas da Política, que são:

- a) a fronteira imperfeita entre jornalistas profissionais e diletantes.
- b) a confusão, muitas vezes proposital, entre textos de opinião e matérias jornalísticas.

c) o alto valor de face que é dado ao que se apresenta notícias de “bastidores”, nas quais as fontes não são citadas, mas critérios que parecem autênticos para proteger fontes (uma pessoa muito próxima ao governador; o “Palácio Anchieta” etc) acabam sendo usados para sustentar a veracidade de informações ou de jogadas escondidas do Público. Só que os atores políticos usam a imprensa para gerar este tipo de impressão e para prejudicar seus adversários; testar reações (balões de ensaio) ou desviar a atenção (chamariz, cortina de fumaça).

d) a espetacularização da notícia, gerando foco preferencial em escândalos de corrupção, ancorados em linguagem moralista (do jornalista ou da suposta fonte), condenatória, que dificulta uma leitura crítica do fenômeno, buscando gerar engajamento de leitores em comentários.

e) uso de recursos digitais nem sempre confiáveis, como trechos de conversas de aplicativos de comunicação, de vídeos, de e-mails, em particular quando são vagos, imprecisos e acabam necessitando de interpretação inserida no texto que os suporta.

f) os interesses dos proprietários de portais de notícias, muitas vezes de caráter regional e temporário, que pode surgir, por exemplo, em pesquisas de opinião de institutos regionais pouco confiáveis.

Além de tudo o que foi dito, levando à necessidade de clareza dos critérios para gerar classificações com base neste tipo de material, certas lacunas tiveram que ser preenchidas com informações colhidas, por exemplo, em postagens em redes sociais. Para o livro, uma solução apresentada com relação à autenticidade foi justamente a apontada por Domingues: a indicação dos *links*. O problema está justamente na permanência dos materiais na internet, o que, a princípio, se espera que seja mais garantido em jornais *online* de empresas estabelecidas, diminuindo o grau de autenticidade por este critério (disponibilidade) para *blogs*, jornais pseudoprofissionais e postagens de redes sociais.

A equipe de pesquisa buscou salvar não só os *links* como também os arquivos em si, mas sem condições de financiamento para disponibilizá-los com o livro, e até mesmo por conta de problemas de *copyright*. Foi gerado, portanto, um acervo privado do projeto, com os *links* salvos, mas sua disponibilidade ao público depende dos seus criadores. A relação dos acervos privados de materiais obtidos *online* para pesquisa, que compõem um tipo específico de repositório destino, com os repositórios fonte é diferente do que foi citado acima.

Quando foi possível e razoável para a pesquisa, buscou-se fazer citações diretas que sustentassem as afirmações, como, aliás, é comum na pesquisa em ciências humanas. Só que, muitas vezes, são necessárias paráfrases, de forma que a confirmação dos dados depende de acesso ao material, que pertence a certas empresas e, mesmo que elas o cedam, implicam custos altos para serem disponibilizados. Desta forma, a pesquisa em questão ilustra um tipo de estudo que se baseia, do ponto de vista dos acervos documentais, em níveis variados de autenticidade, o que se mostra insanável em muitas situações.

Análise do discurso e arquivos digitais

Uma das minhas atividades de docência inclui o ensino de métodos de Análise do Discurso na graduação em História da UFES. Para tanto, são utilizadas publicações do Congresso Nacional, resultante da digitalização dos seus anais. Qual é a especificidade aqui? As turmas devem lidar com séries documentais. Neste sentido, uma das tarefas prescritas (PEREIRA, 2015) é a de, após a definição de um tema a ser abordado na base documental, procurar por todas as suas ocorrências. Um dos objetivos é o de verificar se, quando o assunto era abordado naquele tipo de meio de comunicação, havia padrão discursivo (CARDOSO, 1997), o que os estudantes são estimulados a concluir pela comparação simples dos textos. A partir daí, uma questão relativa à estrutura interna do discurso já pode ser abordada e se torna possível elaborar um critério para a definição do documento, a ser retirado da série, que servirá à análise.

Na graduação, é comum que estudantes tenham dificuldade em definir bons recortes de pesquisa, especialmente no que se refere às datas de início e de fim, de forma a facilitar a pesquisa na base de dados. Além disso, eles tendem a facilitar seu trabalho, procurando por trechos documentais logo no início do período. O ponto a ser destacado aqui e que remete à autenticidade não dos arquivos individuais, mas toda a série documental, é que ela deve ofertar todos os documentos que dela fazem parte. A busca por padrões discursivos (ou a verificação de que eles não existiram no caso em tela) implica o acesso seguro a todas as vezes nas quais as ocorrências da série se deram quando foram produzidas.

Em se tratando do Diário do Congresso Nacional (anais da Câmara dos Deputados) ou dos Anais da República (anais do Senado Federal), o acervo tem que disponibilizar todos os seus números e que estejam em boas condições de leitura. Ora, tanto em repertórios digitais quanto nos físicos, acessados diretamente em bibliotecas ou arquivos, o critério citado nem

sempre é cumprido. **Quando há lacunas na série documental, sua autenticidade diminui.** Neste sentido, seria muito útil se fosse prática corrente e padronizada dos acervos digitais oferecer, de forma acessível a quem consulta, os dados sobre a consistência da série, suas lacunas e seu estado. Se o material foi digitalizado, deveria haver um relatório sobre suas condições originais. Se o material era originalmente digital, informações precisas sobre trechos corrompidos, sobre arquivos ausentes etc. Isto pode até ser exibido por meio de tabelas ou gráficos estatísticos, de forma a fazer uma melhor avaliação do **grau de autenticidade** envolvido.

Perda de arquivos

Ao longo dos anos, atuei em várias pesquisas que demandaram a obtenção de arquivos digitais a partir de repositórios fonte. Quando se trata, por exemplo, de jornais *online*, de plataformas de governos ou sites de instituições, como sindicatos de trabalhadores, existe a possibilidade de que, se o *download* não for feito de forma constante e diária (ou com acessos periódicos para verificar novidades), haja perda de materiais empíricos. O primeiro caso é até mais simples, pois se torna necessário adquirir um hábito. Os jornais *online*, muitas vezes, deixam os *links* das notícias ativos por um prazo determinado, encurtando bastante o tempo disponível para serem baixados.

Já páginas de governos ou instituições passam por mudanças de mandato, que costumam levar a alterações profundas delas. Isto pode resultar na suspensão de conteúdos. Só que estes repositórios fonte costumam ter atualizações aleatórias, pois ocorrem de acordo com circunstâncias específicas, como é o caso de sindicatos de trabalhadores. Eles tendem a publicar periódicos próprios e, eventualmente, boletins específicos, como em épocas de greves ou campanhas salariais. Se não forem baixados a tempo, podem ser perdidos para sempre, até pelo fato de que, mais recentemente, estes materiais tenderam a ser gerados exclusivamente em meio eletrônico.

Outro aspecto que afeta a pesquisa tem a ver com a formação de tais periódicos como série documental. Acima, foram consideradas séries com periodicidade fixa, como jornais *online* ou anais do Legislativo. Só que publicações de sindicatos podem não ter esta característica. Se os prazos de publicação forem largos, quando se estuda uma greve, por exemplo, fatos se desenrolam e acabam não sendo abordados nos documentos. Isto resulta em

perda de informação, o que exige de quem faz o estudo que elabore hipóteses sobre o que houve, situação pela qual já passei (PEREIRA, 2022). É óbvio que este tipo de problema é antigo e ocorre em arquivos físicos.

A novidade aqui é que o progressivo abandono de publicações em papel pelas eletrônicas e a falta de confiança de que permanecerão *online* por tempo indeterminado se soma às lacunas de informações que já fazem parte deste tipo de série produzida de forma espaçada. Mais uma vez, a autenticidade é afetada e varia de acordo com as alterações dos repositórios. De forma geral, havendo pesquisas que envolvam mandatos governamentais, sindicatos e empresas, o trabalho de coleta de arquivos digitais dos repertórios fonte deve ser feito o mais rápido que se possa e guardados em repositório destino privados de pesquisa. Por segurança, não pode haver muita seleção, de forma que estamos falando de uma tarefa que ocupa bastante tempo e pode gerar um volume imprestável (que não será usado) de dados. Ou seja, o que classicamente é chamado de heurística (CARDOSO, 1984), a fase de coleta das fontes, precisa ser estendida por conta da possibilidade concreta de perda de documentos.

Perda de elementos compósitos

Há dificuldade em salvar componentes imagéticos ou sonoros de fontes digitais, em particular quando são incorporados ao texto sem a possibilidades de exibição em separado para armazenamento. Um exemplo neste sentido foi baixado da página do jornal *online* A Gazeta. O título da matéria é “Avenida Civit I, na Serra, vai ser ampliada e ganhar parque; veja vídeo”. É de autoria da redação do jornal. O *link* é: <https://www.agazeta.com.br/es/cotidiano/av-civit-i-na-serra-vai-ser-ampliada-e-ganhar-parque-veja-video-0923>. Foi baixado em 1 de setembro de 2023. A matéria é acompanhada por uma animação em computação gráfica, que foi enviada ao periódico pela Prefeitura da cidade da Serra, no Espírito Santo (ES).

Além do vídeo, há quatro imagens que constam de fotogramas do vídeo em questão. Na história recente do ES, a apresentação de imagens geradas por computação com o objetivo de materializar promessas, especialmente relacionadas ao setor de transportes ou construções de empresas, são muito comuns. Mais comum ainda é o fato de tais promessas não serem cumpridas. Ou seja, as imagens em questão são componentes iconográficos de um discurso no qual estes se encontram em diálogo direto com os elementos textuais. O mais relevante é o aspecto assertivo desta interação. O primeiro parágrafo da referida matéria diz:

A Avenida Civit I, no município da Serra, na Região Metropolitana de Vitória, vai ser ampliada. De acordo com projeto da prefeitura, apresentado na última quarta-feira (30), o trecho entre o trevo de Maringá e a BR 101 passará a ter três faixas em cada sentido, separadas por um canteiro central. Atualmente, a via conta com apenas uma faixa em cada sentido (AVENIDA, 2023).

O texto afirma que a avenida “vai ser ampliada”. Na sequência é que surgem os condicionantes, como: “a previsão é de que a obra comece em dezembro deste ano, a depender da licitação”. Ou seja, nada está garantido; o tom geral do discurso é o de que já está tudo decidido. A narração do vídeo afirma: “[...] o projeto também pretende atrair novos empreendimentos para a região e, com isso, criar novos empregos”. A referida promessa não se encontra no texto. O projeto inclui um campo de futebol e outros equipamentos urbanos, com parque, pistas de caminhada e skate. Um dos quatro fotogramas que acompanham o texto é mostrado a seguir:

Figura 1 – Fotograma da matéria



Fonte: <https://www.agazeta.com.br/a7b359e2-e6b9-4322-a33c-e694b3a93bf4>.

Como é bem típico deste tipo de produção pictórica, pessoas são ausentes ou mostradas à distância. Não há sujeira. Carros e galpões de empresas se destacam, além da obra com belas faixas de rodagem. O discurso aí presente toma uma promessa de futuro como

realidade já decidida e configurada de acordo com expectativas estéticas mais relacionadas ao mundo dos *games* virtuais do que da realidade. O exemplo que está sendo usado aqui é apenas uma ilustração de algo absolutamente comum ao longo dos últimos anos no estado do ES. Muitas promessas deste tipo foram feitas e não foram cumpridas. Trata-se de uma fala que tematiza o desenvolvimento, o crescimento como uma esperança, particularmente no que se refere à criação de empregos.

Os setores dominantes da sociedade capixaba elaboraram um projeto de poder que se sustenta na ideia de que o maior problema do estado seriam os chamados “gargalos logísticos”, ou seja, as supostas ineficiências dos transportes: rodovias, vias urbanas, portos, aeroportos, ferrovias etc. Qualquer medida apresentada como intervenção de peso no setor é sempre bem-vinda pela imprensa, que faz o tratamento no mesmo sentido do que foi apresentado aqui: obras materializadas em computação gráfica e associadas à criação de empregos.

Qual é a autenticidade de um arquivo eletrônico como este, que seja salvo com o texto, o vídeo e os fotogramas? Ele não só não reflete a realidade empírica vigente, o que fica bem evidenciado, pois a obra é futura, como reproduz um discurso de poder, seja das elites políticas, seja da imprensa. A mídia é sócia deste projeto, na medida em que ele oferece meios para gerar notícias em tom positivo, alimentando esperanças e, ao mesmo tempo, combatendo representações negativas do mundo. A cidade da Serra, no caso, é um local estigmatizado pelo noticiário policial, que também é um dos principais componentes da cobertura jornalística.

Neste sentido, a grande imprensa executa um trabalho de “morde e assopra”: os aspectos positivos e negativos são pesados, distribuídos na abordagem da realidade por meio das editorias. A de “cidades”, “economia” ou “política” podem abrir espaço para promessas que compensem as duras notícias e imagens da cobertura policial. Ou seja, as múltiplas abordagens, configuradas em meios que são transformados em arquivos digitais, precisam ser consideradas como séries documentais que constituem um todo.

Há uma série bem definida de promessas de desenvolvimento com todos os aspectos citados acima, ao longo dos últimos anos. Ela não deve ser considerada de forma isolada, pois funciona como contraponto ao enorme volume de más notícias, que tendem, inclusive, a serem mais espetaculares, pois refletem bem mais a sensação de (in)segurança, intensificada

pelo formato das publicações. Logo, os conjuntos de séries contêm discursos sobre o real que dialogam entre si de forma indireta. E, ao mesmo tempo, sustentam projetos de poder.

É óbvio que uma obrigação de quem pesquisa é conseguir perceber este conjunto e fazer sua análise como um todo. A questão aqui é a relevância de que todas as séries que dialogam indiretamente entre si estejam disponíveis e possam ser baixadas, incluindo os elementos compósitos, como os iconográficos, pois, digamos, a partir deste exemplo, pares de imagens, positivas para as obras, negativas para a criminalidade, são componentes do discurso como um todo. Como dito antes, documentos que são publicados na imprensa em arquivos de imagens/áudio/vídeo nem sempre podem ser salvos e, além disso, ocupam muito espaço de armazenamento.

Variabilidade

No caso de periódicos de notícias, ao longo do tempo, foram desenvolvidas “boas práticas” de correção de informações. O padrão passou a ser a indicação dos acertos em edições seguintes. Alguns diários, de forma mais responsável, criaram a chamada “Erramos” para destacar o ponto. Outros o faziam em partes menos visíveis de suas páginas. Na atualidade, contudo, a prática foi alterada, pois é costume deixar a matéria de jornal *online* aberta para correções e atualizações. Quando elas são salvas antes das alterações, surge um problema. Isto remete à observação de Domingues que envolve “variabilidade limitada, considerando-se [...] uma variação de forma e do conteúdo do documento que não compromete seu caráter arquivístico à medida que é implementada por regras fixas, o que equivale a dizer que tal variação é intencionada pelo autor” (DOMINGUES, 2023, p. 22, citando Rondinelli). Neste sentido, as mudanças não capturadas afetam ao critério da variabilidade limitada, que é uma das questões envolvendo autenticidade.

A prática anterior, contudo, foi mantida no caso de arquivos *online* fechados, como textos em pdf, de forma que correções posteriores, ainda podem ser publicadas em números seguintes dos jornais. E isto pode ser combinado com, por um lado, textos fechados em pdf e, por outro lado, textos dinâmicos em *links*, sobre o mesmo assunto, que podem ser modificados. Tudo isso tem a ver com a pressa dos jornais *online*, no formato de portais de notícias, em publicar rapidamente para “dar furo” nos outros meios de comunicação, que intensificou este tipo de problema. A pressa, sem a devida checagem de fatos ou oferecimento

de outra versão, também deriva do lugar ocupado pelas redes sociais na geração de “notícias”. Aqui, portanto, há uma questão correlata ao item anterior, que é necessidade de investimento heurístico na coleta específica de correções.

Conclusões

Usar arquivos originalmente digitais para a pesquisa em história do tempo presente envolve muitos tipos de dificuldades, que vão, inclusive, mudando com o tempo. Mais recentemente, temos as variadas formas automatizadas de elaboração de narrativas, como *fakes* e *bots* pré-programados, além de aplicativos com recursos de inteligência artificial. Isto afeta a identificação da autoria. Há pessoas reais usando programas que geram conteúdos apresentados como se fossem legitimamente seus, a exemplo de geradores de lero-lero (RAMOS, 2012) ou, mais recentemente, o *ChatGPT* (DENG; LIN, 2022) e seus congêneres. Há pessoas que usam recursos *online* para falsear autoria de discursos, como no caso dos *deepfakes* (FALLIS, 2020). E há sistemas totalmente automatizados fazendo coisas semelhantes.

É claro que a pesquisa em História sempre lidou com falsidades, a ponto deste elemento ter sido incorporado por Langlois e Seignobos (1992) em seu célebre manual, originalmente publicado em 1898, por meio dos procedimentos específicos de crítica externa. A questão que desejo destacar e que é correlata à discussão sobre autenticidade que estou fazendo aqui é que **o universo dos arquivos digitais gera dificuldades novas**, como a preocupação constante em baixar os arquivos, de forma atualizada e a intensificação do problema da autoria. Bastante trabalho acaba tendo que ser despendido em etapas anteriores à análise de conteúdo em si. As fronteiras entre documentos com autoria e conteúdos confiáveis não só diminui como o seu volume aumenta consideravelmente. Eles são autênticos como mentiras elaboradas de propósito, fazendo parte estrutural dos ecossistemas de desinformação (SANTANA, 2023). A questão é classificá-los.

Acima, falei bastante dos arquivos de curto prazo. As pesquisas recentes sobre o fenômeno do bolsonarismo, por exemplo, colhem dados de grupos de aplicativos de comunicação por celulares, como o *WhatsApp* e o *Telegram*, nas redes sociais, em fóruns e na *deepweb*. Muitos destes conteúdos foram feitos de propósitos para publicação por pouco tempo. Alguns foram retirados do ar por decisão judicial e a autoria é dissimulada. Tais pesquisas formam repertórios destinos privados que, ao final, não poderão ser comparados

com suas fontes originais. Um caminho para o futuro será a recuperação destes acervos privados e sua disponibilização. Contudo, eles poderão carregar consigo várias das limitações que fui citando ao longo deste trabalho.

À guisa de conclusão, portanto, os pontos principais que sugeri para ampliar a discussão sobre o tema foram os seguintes:

a) Devemos distinguir arquivos com elementos que foram digitalizados e arquivos que foram criados originalmente como digitais.

b) Acervos contendo arquivos digitais individuais ou séries documentais digitais podem apresentar graus variáveis de autenticidade.

c) O conceito de autenticidade dos arquivos digitais deve ser concebido em um **contínuum**, de um maior a menor grau.

d) Devemos considerar a diferença entre repositório fonte e destino.

e) Muitos acervos de arquivos digitais não foram feitos para serem oferecidos por tempo indefinido.

f) Arquivos digitais costumam possuir caráter compósito, com elementos temporários, o que dificulta o seu *download* completo.

g) Os custos de armazenamento e de heurística deste tipo de documento não devem ser desprezados, afetando a sua existência.

h) Há problemas de variabilidade dos arquivos digitais.

O tema é de grande relevância para a pesquisa, remetendo a sérios problemas metodológicos que não devem ser ignorados.

Referências

AVENIDA Civit I, na Serra, vai ser ampliada e ganhar parque; veja vídeo. **A Gazeta**, Vitória, 1 set. 2023. Disponível em: <https://www.agazeta.com.br/es/cotidiano/av-civit-i-na-serra-vai-ser-ampliada-e-ganhar-parque-veja-video-0923>. Acesso em: 1 set. 2023.

BRUNI, José C.; MENNA-BARRETO, Luiz; MARQUES, Nelson (Orgs). **Decifrando o tempo presente**. São Paulo: Editora UNESP, 2007.

CARDOSO, Ciro F. S. **Uma introdução à História**. 4. ed. São Paulo: Brasiliense, 1984.

CARDOSO, Ciro F. S. **Narrativa, sentido, história**. Campinas: Papiрус, 1997.

COSTA, Israel R.; PINTO, Liliane F. C. **A evolução dos dispositivos de armazenamento de dados na perspectiva da história.** [2017]. Monografia (Graduação em Informática) – Universidade Federal do Maranhão, Codó (MA), [2017]. Disponível em: <https://rosario.ufma.br/jspui/bitstream/123456789/2830/1/IsraelReisCosta.pdf>. Acesso em: 19 set, 2023.

DENG, Jianyang; LIN, Yijia. The benefits and challenges of ChatGPT: an overview. **Frontiers in Computing and Intelligent Systems**, v. 2, n. 2, p. 81-82. Disponível em: <https://drpress.org/ojs/index.php/fcis/article/view/4465>. Acesso em: 26 set. 2023.

DOMINGUES, Jonathan M. Autenticidade dos arquivos digitais nas pesquisas históricas da educação matemática. **Convergências: Estudos em Humanidades Digitais**, Goiânia, v. 1, n. 1, p. 16-31, jan/abr 2023. Disponível em: <https://periodicos.ifg.edu.br/cehd/article/view/78/73>. Acesso em 26 set. 2023.

LANGLOIS, Chares-Victor; SEIGNOBOS, Charles. **Introduction aux études historiques.** Paris: Kimé, 1992.

FALLIS, Don. The epistemic threat of Deepfakes. **Philosophy & Technology**, v. 34, p. 623–643, 2021. Disponível em: <https://link.springer.com/article/10.1007/s13347-020-00419-2>. Acesso em: 21 set. 2023.

PEREIRA, André R. V. V. **Teoria e metodologia da história:** o tratamento de fontes. Vitória: UFES/SEAD, 2015.

PEREIRA, André R. V. V.. A Samarco, os trabalhadores e o poder público: uma coalizão produtivista como fundamento do rompimento da barragem de Fundão e suas consequências. In: LAVALLE, Adrian G; CARLOS, Euzeneia (Orgs). **Desastre e governança no Rio Doce:** atingidos, instituições e ação coletiva. Rio de Janeiro: Garamond, 2022. p. 81-105.

PEREIRA, André R. V. V.; ZAIDAN, Junia C. S. M.; GALVÃO, Anda C. **A invenção da balbúrdia:** dossiê sobre as intervenções de Bolsonaro nas instituições federais de ensino superior. Brasília: ANDES, 2022.

RAMOS, François S. **Fraude acadêmica:** uma análise ético-legislativa. 2012. Dissertação (Mestrado em Educação) - Programa de Pós-Graduação em Educação da Universidade de Uberaba, Uberaba, 2012. Disponível em: <https://repositorio.uniube.br/bitstream/123456789/787/1/FRAN%c3%87OIS%20SILVA%20RAMOS.pdf>. Acesso em: 27 set. 2023.

REIS, Arthur F. As Humanidades Digitais no Brasil e no mundo: o estado da arte. **Convergências: Estudos em Humanidades Digitais**, Goiânia, v. 1, n. 1, p. 32-48, jan/abr 2023. Disponível em: <https://periodicos.ifg.edu.br/cehd/article/view/47/74>. Acesso em: 26 set. 2023.

SANTANA, Eliara. Ecosistema de desinformação se consolidou com o bolsonarismo. In: AVRITZER, Leonardo; SANTANA, Eliara; BRAGATTO, Rachel C. (Orgs). **Eleições de 2022 e a reconstrução da democracia no Brasil.** Belo Horizonte: Autêntica, 2023. p. 89-101.



SCHUINA, Lucas. Novo portal reúne material sobre patrimônio histórico e cultural do estado. **Século Diário**, Vitória, 26 set. 2023. Disponível em: <https://www.seculodiario.com.br/cultura/novo-portal-reune-material-sobre-patrimonio-historico-e-cultural-do-estado>. Acesso em: 26 set. 2023.

SOUZA, Rildo B. de. Possibilidades de pesquisa em acervos digitais: os relatos de viagens sobre Goiás do século XIX. **Convergências: Estudos em Humanidades Digitais**, Goiânia, v. 1, n. 1, p. 1-15, jan/abr 2023. Disponível em: <https://periodicos.ifg.edu.br/cehd/article/view/106/72>. Acesso em: 25 set. 2023.

Recebido em: 30 de setembro de 2023

Aceito em: 3 de novembro de 2023
