

Arquivo, memória e Big Data: uma proposta a partir da Covid-19

*Ian Kisil Marino^I
Pedro Telles da Silveira^{II}
Thiago Lima Nicodemo^{III}*

Resumo: Este artigo busca refletir sobre os impactos das tecnologias digitais nas formas de arquivamento contemporâneo e apresentar a perspectiva de atuação elaborada no âmbito do projeto Memória Covid-19 Brasil, através do trabalho conjunto dos integrantes do Centro de Humanidades Digitais IFCH-UNICAMP e do projeto DéjàVu, da mesma instituição. Primeiro, abordaremos o caráter da memória na contemporaneidade, as tecnologias digitais e o deslocamento do lugar do historiador, principalmente tendo em vista o desafio apresentado pela pandemia do novo coronavírus. Depois, faremos um breve levantamento das iniciativas de arquivamento e constituição da memória da Covid-19 no país, para, em seguida, ressaltar as especificidades da proposta do Centro de Humanidades Digitais IFCH-UNICAMP, incluindo a criação de ferramentas informatizadas para a coleta de publicações. Por fim, retomaremos a importância da adoção de uma postura mais ativa pelos historiadores diante dos problemas contemporâneos, o que implica, também, repensar a constituição das nossas equipes de pesquisa e o aprofundamento da interdisciplinaridade de nossa atuação.

Palavras-chave: História Digital; Memória; Covid-19; Arquivos Digitais; Teoria da História.

Archive, memory, and Big Data: a proposal based on Covid-19

Abstract: This article seeks to reflect on the impacts of digital technologies on the contemporary archiving methods and to present the perspective of action prepared under the *Memória Covid-19 Brasil* [Covid-19 Memory Brazil] project, through the joint work of the members of *Centro de Humanidades Digitais IFCH-UNICAMP* [Digital Humanities Center – IFCH-UNICAMP] and of the *Déjà Vu* project, from the same institution. First, it approaches the character of memory in contemporary times, the digital technologies, and the displacement of the historian's work, especially regarding the challenge imposed by the new coronavirus disease. Then, we will briefly raise the initiatives of archiving and constructing the Covid-19 memory in the country, so that afterward we may highlight the specificities of the proposal by *Centro de Humanidades Digitais IFCH-UNICAMP*, including the creation of computer-based tools for the collection of publications. Lastly, we will recall the importance of the adoption of a more active attitude by historians in the face of contemporary issues, which also implies to reconsider the constitution of our research groups and the further development of our practice's interdisciplinarity.

Keywords: Digital History; Memory; Covid-19; Digital Archives; Theory of History.

Artigo recebido em 20/07/2020 e aprovado em 25/07/2020.

ARQUIVO, MEMÓRIA E BIG DATA: UMA PROPOSTA A PARTIR DA COVID-19

IAN KISIL MARINO
PEDRO TELLES DA SILVEIRA
THIAGO LIMA NICODEMO

Em 12 de novembro de 2019, o jornal inglês *The Guardian* publicou uma extensa matéria sobre o trabalho do biólogo evolucionista e especialista em políticas públicas Peter Turchin. Para além do seu campo de especialização, Turchin também voltara sua atenção para a história, buscando apontar, probabilisticamente, a correlação de fatores responsáveis por fazer as sociedades entrarem em crise. O trabalho de Turchin é tanto ingênuo quanto consciente das limitações inerentes ao papel dos dados quantitativos na análise histórica: ao ser perguntado sobre o reforço de vieses interpretativos pelos dados, utilizados em qualquer amostragem analisada, ele afirma que todas as interpretações são parciais, mas aquelas que possuem conjuntos maiores de dados, ao menos, têm meios para corrigir as suas tendências negativas.

A postura de Turchin é menos interessante pelos resultados aos quais possa chegar do que pelos aspectos metodológicos que a condicionam. Sua análise renova a esperança na descoberta de padrões e recorrências no processo histórico que permitam antever e prevenir acontecimentos catastróficos. Entretanto, ela também se beneficia de uma paisagem de dados e de um esforço de quantificação que atravessou as mais variadas disciplinas científicas, incluindo a história, ao longo dos últimos cinquenta anos, para enfim chegar ao mercado no início do século XXI. Estamos falando, é claro, do Big Data.

As grandes bases de dados são alvo de reações contraditórias dos historiadores. Por um lado, a emergência do Big Data é vista com suspeita por motivos tanto políticos – o caráter proprietário das principais plataformas digitais e a concentração de renda e poder que lhe segue, para não mencionar as desigualdades de acesso às tecnologias digitais – quanto epistemológicos – a promessa de que os dados substituirão a tecnologia e o olhar humano, tornando o próprio historiador obsoleto. Por outro lado, é difícil não imaginar que um pesquisador não celebraria possuir mais dados, que tornassem menos fragmentado o registro histórico dos períodos que estuda. A tecnofilia e a tecnofobia são sintomas de uma relação problemática com a técnica, já apontou Gilbert Simondon^{IV}; mesmo assim, apenas Jo Guldi e David Armitage^V, entre os historiadores com maiores credenciais na disciplina histórica, advogaram usos criativos do Big Data em anos recentes.

O problema muda de figura, no entanto, se passarmos do presente para o futuro das pesquisas históricas. Mais especificamente, o problema das fontes que os historiadores terão para analisar o período que vivemos. Roy Rosenzweig^{VI}, ainda nos anos 1990, já alertara que a fronteira entre a escassez e a abundância documental, proporcionada pelas mídias digitais, é bastante tênue. Desde então, esta sensação não fez senão crescer. Mesmo assim, mais de vinte anos após o historiador norte-americano ter soado o seu alerta, talvez estejamos chegando num ponto em que os historiadores obtenham maior autonomia frente a oscilação entre a escassez e a abundância. Para isso, o próprio Big Data pode nos ajudar.

Este artigo busca refletir sobre os impactos das tecnologias digitais sobre as formas de arquivamento contemporâneo e apresentar a perspectiva de atuação elaborada no âmbito do projeto Memória Covid-19 Brasil, em especial através do trabalho conjunto dos integrantes do Centro de Humanidades Digitais IFCH-UNICAMP e o projeto DéjàVu, na mesma instituição. Em primeiro lugar, abordaremos o caráter da memória na contemporaneidade, as novas tecnologias e o deslocamento do lugar do historiador, principalmente tendo em vista o desafio apresentado pela pandemia do novo

coronavírus; depois, faremos um breve levantamento das iniciativas de arquivamento e constituição da memória da Covid-19 no país, para, em seguida, ressaltar as especificidades da proposta do Centro de Humanidades Digitais IFCH-UNICAMP, incluindo a criação de ferramentas informatizadas para a coleta de publicações. Por fim, retomaremos a importância da adoção de uma postura mais ativa pelos historiadores diante dos problemas contemporâneos, o que implica, também, repensar a constituição das nossas equipes de pesquisa e o aprofundamento da interdisciplinaridade de nossa atuação.

Os arquivos digitais e a história

A intensa popularização das tecnologias digitais – ocorrida, genericamente, do estabelecimento do computador pessoal à massificação do uso da internet, já no século XXI – trouxe novos parâmetros e práticas para as relações interpessoais, profissionais e disciplinares. Não faltam à história nem evidências e nem pesquisadores que apontem para a já cotidiana presença das mídias digitais no universo de trabalho dos historiadores. O ritmo acelerado de pesquisa e divulgação, a interação com fontes intermediadas pelo processamento de dados e pela chamada inteligência artificial, a dependência de servidores de rede e de *softwares* e ferramentas de trabalho privados, entre outros, ilustram a cena do historiador contemporâneo, imerso em um panorama muito diferente daquele em que foram formulados os fundamentos teóricos e metodológicos da disciplina histórica^{VII}.

Se a presença do digital é inequívoca, a compreensão dos seus efeitos sobre a constituição epistemológica da história suscita inúmeras questões, que ainda precisam ser respondidas. Retomando a constituição disciplinar da área, percebe-se que a ocorrência do digital requer esforços de reflexão teórica sobre o que é a historiografia no atual cenário. Vale lembrar que a história se consolidou, no século XIX, no esteio do estabelecimento estrutural e burocrático dos estados modernos europeus, no ascender das narrativas nacionais e no alargamento das fronteiras comerciais e produtivas do capitalismo em ascensão. Os parâmetros epistemológicos da disciplina histórica não foram exatamente determinados por esses eventos, mas foram concebidos através da interação orgânica e como parte desses processos^{VIII}. Não seria possível imaginar, portanto, que o cenário atual, marcado pela universalização das tecnologias digitais, pela desestruturação dos estados nacionais frente a grandes corporações proprietárias das ferramentas de processamento e armazenamento de dados, contribuiria com a configuração de uma disciplina histórica diversa?

Dentre todo o possível campo de atuação da história digital, esta contribuição elege o arquivo como objeto central. Procura-se entendê-lo segundo duas vertentes, dialógicas e simultâneas. Uma, mais prática, entende o arquivo como a instância responsável pela preservação dos materiais diversos para usos – também diversos – futuros. Essa leitura compreende o arquivo como uma instituição, destacando a sua constituição física de espaço de armazenamento de documentos. Ela compreende, naturalmente, uma variedade de instâncias, como arquivos temporários circulantes, ligados à burocracia estatal; arquivos históricos nacionais públicos; arquivos pessoais; arquivos digitalizados e nascidos digitais. Uma segunda perspectiva compreende o arquivo como uma prática. Arquivar é deliberar sobre o lembrar e o esquecer, é a manifestação da autoridade sobre o legado da experiência humana, a partir da

ARQUIVO, MEMÓRIA E BIG DATA: UMA PROPOSTA A PARTIR DA COVID-19

IAN KISIL MARINO
PEDRO TELLES DA SILVEIRA
THIAGO LIMA NICODEMO

manipulação, preservativa ou destrutiva, dos seus vestígios. Nesse sentido, discutido sobretudo por Michel Foucault e Jacques Derrida, o arquivo é um dispositivo de poder, incluindo o seu agenciamento sobre as formas de representação do passado, desdobradas nas relações interpessoais do presente e capacitando as projeções para o horizonte futuro^{IX}. Essas duas concepções sobre a noção de arquivo justificam o interesse teórico da história sobre ele: o arquivo se configura como uma instância de modulação de memórias em discursos narrativos imaginativos sobre o passado, função que é análoga à da escrita da história.

Se a história é constituída a partir de suas fontes, que lhe servem de evidências, o arquivos é uma espécie de “pré-história”, como caracterizou Andreas Fickers, na medida em que nele se fizeram as escolhas preliminares sobre a preservação ou o descarte de determinados objetos – fontes históricas em potencial^X. Situado num momento anterior à pesquisa histórica, o arquivo configura um processo de recorte de evidências, de seleção de vestígios para a posteridade, o que influi profundamente no trabalho dos historiadores. Assim, a compreensão e o mapeamento do processo arquivístico, como proveniência moduladora de fontes históricas, são tópicos da reflexão historiográfica. Ou, como já ressaltou Mario Wimmer^{XI}, os arquivos são um problema teórico da historiografia.

Desde a década de 1980, iniciou-se um movimento historiográfico de aproximação com o estudo teórico dos arquivos, motivados tanto pela interpretação foucaultiana quanto pela ascensão das correntes como o New Historicism, a nova história política e a histórica cultural francesas, a microhistória italiana, as diferentes vertentes da biografia histórica, a história intelectual na base proposta pela Escola de Cambridge e os estudos pós-coloniais^{XII}. Compreendendo a influência vinculante e sistemática do arquivo sobre a pesquisa histórica, assim como a partir da constatação dos limites dos arquivos públicos, esses esforços multiplicaram as formas de arquivamento, além de inspirarem as reflexões que, por fim, culminaram no esforço de compreender, epistemologicamente, o arquivamento digital. Essa missão vem se mostrando complexa, na medida em que, no conjunto das mudanças técnicas introduzidas pelas tecnologias digitais, a proveniência e os critérios arquivísticos se tornaram fatores nebulosos^{XIII}.

A designação “arquivo digital” emergiu entre o final da década de 1990 e o início dos anos 2000, referindo-se inicialmente a coleções armazenadas na internet. Desde então, a categoria se diversificou, passando a abranger tanto os acervos digitalizados, abrigados em instituições de renome, quanto as experiências informais e efêmeras de arquivos nascidos digitalmente. Todas as experiências, no entanto, apresentam o mesmo desafio, qual seja, a necessidade de observar seus processos de arquivamento com olhar historiográfico. Aludindo aos meandros matemáticos que projetam, na tela do computador, um documento de arquivo digital, e reforçando a demanda pelo estudo crítico dos arquivos digitais, Ina Blom^{XIV}, comenta que “com as imagens digitais, nada é armazenado senão o código: o próprio potencial para gerar uma imagem de um certo compósito material de novo e de novo por meio de constelações numéricas”. A decodificação de dados comprimidos em linguagem numérica, que dispara as ações pré-programadas para que o arquivo disponibilize o documento buscado, podem até se assemelhar aos recursos catalográficos da organização de um arquivo tradicional. Entretanto, essa característica sublima, ao menos, duas diferenças

ARQUIVO, MEMÓRIA E BIG DATA: UMA PROPOSTA A PARTIR DA COVID-19

IAN KISIL MARINO
PEDRO TELLES DA SILVEIRA
THIAGO LIMA NICODEMO

centrais em relação ao arquivamento analógico: a composição do documento e o agenciamento sobre o processo arquivístico. Como destaca Michael Moss^{XV},

quando vemos um documento online, seja uma imagem, um texto ou uma peça musical, estamos olhando para um objeto mantido num servidor por um provedor. O provedor pode ter cópias adicionais por motivos práticos ou de segurança, mas o acesso é muito diferente do manuscrito analógico e da cultura impressa.

O documento digital – ou mesmo o digitalizado –, quando visível, é uma projeção de operações codificadas, diferente do documento material de um arquivo analógico, que se apresenta em si mesmo. As operações de projeção desse documento, ainda, não são uma ação direta do curador ou arquivista, mas parte de um protocolo de decodificação digital pré-programado. Esse processo ocorre de forma mais ou menos autônoma, na mesma medida em que é indiferente à agência de uma pessoa ou ao conteúdo do documento codificado – o que apenas reforça a necessidade da presença crítica humana para a sua compreensão.

A obscuridade dos processos de arquivamento digital, sobretudo para aqueles não especializados em processamento de dados, como a maioria dos historiadores e humanistas, é patente. O itinerário histórico desenvolvido pelos arquivos, pautados no mapeamento das proveniências e nas escolhas sobre o lembrar e o esquecer parecem ideias fora de lugar no panorama digital. A teoria da história, interessada nos fundamentos arquivístico, também se vê deslocada desse ambiente – o que explica as crises disciplinares e profissionais vividas por arquivistas e historiadores na atualidade.

Os arquivos digitais, parece, dispensam os historiadores e os arquivistas, conseguindo, com a sua própria linguagem e relativa autonomia, tratar os documentos e preservá-los. Essa impressão, entretanto, é não apenas superficial como também sintomática de problemas mais sérios, como o definhamento da memória como espaço público, o controle dos dados pessoais por grandes corporações privadas e a retração dos critérios sobre o lembrar e o esquecer, entregues aos protocolos digitais formulados por agentes desinteressados e profissionalmente incapacitados para o trato público dos vestígios da experiência humana^{XVI}. O arquivo digital é um campo com urgente necessidade de intervenção de profissionais capacitados, como historiadores, não para assumirem postura tecnofóbica frente ao digital, mas para garantirem que se cultivem práticas criteriosas e significativas de arquivamento, demanda que é reforçada pela imensa quantidade de dados em circulação^{XVII}. Apesar de ser intuitivo e inclusivo em várias de suas facetas, o ambiente digital tende a repelir a presença de profissionais não especializados em seu meio, com conceitos e protocolos particulares e inférteis em outras áreas do conhecimento, como apontou Evgeny Morozov^{XVIII}. Os historiadores, no entanto, devem insistir para entrar nesse meio, porque se fazem necessários – como sempre foram – para a compreensão desses processos.

Embora exista uma notável diversidade de projetos e categorias sob o guarda-chuva do “arquivo digital”, as que mais nos interessam são aquelas iniciativas relacionadas à experiência de acontecimentos históricos de grande magnitude e caráter disruptivo. A partir de determinada anormalidade cotidiana, socialmente vivida e perceptível, e do registro individual do ocorrido, desenha-se uma iniciativa de reunião desses registros numa coletânea informal digital, que acaba formando um arquivo. Esse tipo de iniciativa se tornou mais comum na sequência do 11 de Setembro, no Estados

ARQUIVO, MEMÓRIA E BIG DATA: UMA PROPOSTA A PARTIR DA COVID-19

IAN KISIL MARINO
PEDRO TELLES DA SILVEIRA
THIAGO LIMA NICODEMO

Unidos, quando motivou a criação do *September 11th Digital Archive*, iniciativa de grandes dimensões dedicada à coleta e à preservação das memórias de indivíduos comuns daqueles acontecimentos, além de dezenas de outras iniciativas menores^{XIX}. Estes casos têm grande engajamento comunitário, constituindo acervos colaborativos e lugares de memória para eventos imprevisíveis, o que não tem precedentes em tamanha escala sem a presença das tecnologias digitais. Esse tipo de arquivo, criado a partir de colaborações voluntárias digitais, tem sido chamado de *crowdsourcing*, e é bastante valorizado por historiadores identificados com a história pública, corrente preocupada com a ressignificação social e participativa da disciplina histórica.

A experiência da Covid-19 pode ser considerada um desses eventos de grande magnitude e caráter disruptivo. Após a classificação, pela Organização Mundial da Saúde, em 11 de março de 2020, do avanço da doença como uma pandemia, praticamente o mundo todo se viu compelido a mudanças drásticas no padrão de vida, alterando substancialmente as relações familiares, sociais e profissionais. A adoção de políticas de distanciamento social, o isolamento dos infectados pelo novo coronavírus, a restrição do funcionamento econômico dos setores básicos da produção, a incerteza sobre o tratamento da doença e o pânico decorrente dos números de infectados e mortos, para mencionar apenas alguns tópicos, têm marcado o cotidiano da pandemia. O efeito conjunto desses acontecimentos tem alterado as relações sociais em escala global, processo que é acelerado pelo intenso afluxo de informações, formas de trabalho remoto e recursos de entretenimento mediado pelas tecnologias digitais. Ao mesmo tempo, intensifica-se a desigualdade social e econômica em níveis globais, regionais e nacionais, o que tensiona as linhas de força no embate sobre o encaminhamento da crise sanitária e econômica.

A pandemia da Covid-19 vem impactando a sociedade com movimentos de sincronia e assincronia globais, uniformizando experiências no cotidiano do isolamento, no qual o digital se torna, praticamente, o único meio de comunicação e relação social, enquanto outras vozes se silenciam devido às mortes e à diminuição das margens de atuação de sujeitos marginalizados. A relevância disruptiva desse evento justifica o esforço de se contemplar quais serão as formas de mobilização de recursos para a criação de ferramentas para a preservação das diferentes visões desse cenário. Em se tratando de um evento trágico e marcante de escala global, quais os meios para que se criem arquivos para a preservação do legado dessa experiência humana e para seu uso futuro por historiadores, além de membros da comunidade em geral interessados na compreensão dessa experiência?

A pandemia e os arquivos digitais

A pandemia do novo coronavírus pode justamente ser considerada um evento disruptivo em escala global. Como outros acontecimentos recentes de grande envergadura – podemos pensar nos atentados terroristas do 11 de Setembro, no Estados Unidos –, a pandemia impactou todos; ao contrário daquele acontecimento, porém, no caso da pandemia somos todos vítimas desse evento global, agentes de sua realização, e não apenas suas testemunhas^{XX}. Estes eventos disruptivos globais se inserem numa transformação sensível dos significados atribuídos ao histórico nas últimas décadas, já anunciada por Pierre Nora nos anos 1970, para a qual o histórico se refere àqueles acontecimentos imprevisíveis, de grande magnitude e intenso alcance midiático^{XXI}.

ARQUIVO, MEMÓRIA E BIG DATA: UMA PROPOSTA A PARTIR DA COVID-19

IAN KISIL MARINO
PEDRO TELLES DA SILVEIRA
THIAGO LIMA NICODEMO

O retorno do acontecimento através do alcance midiático é uma das formas de recuperação do tempo breve, já assinalada negativamente por Jo Guldi e David Armitage, citados anteriormente. Para os autores, o tempo breve é antitético ao Big Data, que permite justamente estender a escala de análise para intervalos temporais cada vez maiores. Entretanto, o Big Data não está relacionado somente ao tamanho – ao volume de informações reunidas –, mas também à interrelação entre os dados e às inferências que permite fazer. O Big Data também diz respeito ao tempo breve dos acontecimentos – aliás, é onde ele mais propriamente pode ser aplicado ao conhecimento histórico.

Na verdade, os historiadores têm desenvolvido formas cada vez mais sofisticadas de preservar a memória desses acontecimentos disruptivos globais, ainda que poucas vezes através da mobilização ou do auxílio do Big Data. Já em 2005, após a experiência na criação do *September 11th Digital Archive*, os pesquisadores associados ao Roy Rosenzweig Center for History and New Media (CHNM) embarcaram na criação de um segundo arquivo dedicado à captura da memória gestada “ao vivo” relacionada a um desses eventos. Motivados pela passagem dos furacões Katrina e Rita pelo Sul dos Estados Unidos, os pesquisadores criaram o *Hurricane Digital Memory Bank*, em parceria com a Universidade de New Orleans. O projeto tinha como objetivo coletar o máximo possível de materiais relacionados à passagem dos furacões e seu impacto nas comunidades afetadas. Como destacam seus criadores, “assim que o furacão Katrina atingiu o litoral no dia 29 de agosto de 2005, a equipe do CHNM rapidamente percebeu que se estava vivenciando um momento muito significativo da história americana”, de modo que era preciso agir rápido para,

(...) coletar e preservar o máximo da “história instantânea” destes eventos – história que estava sendo criada e publicada por milhares de pessoas comuns em seus blogs pessoais, em serviços de compartilhamento de fotos e no YouTube^{XXII}

O *Hurricane Digital Memory Bank*, portanto, estabelecia uma relação direta com as plataformas de mídia que já caracterizavam a crescente complexidade das tecnologias digitais em suas interrelações com os processos sociais. Mais do que este aspecto, no entanto, destaca-se a possibilidade de construir um arquivo em tempo real, diminuindo a distância temporal que transformava os documentos em vestígios históricos^{XXIII}.

Entretanto, a maior parte dessas iniciativas, incluindo aquelas do CHNM, procura constituir seus arquivos através da colaboração com o público, somando seus esforços por meio do envio de testemunhos relacionados ao evento em questão. Essas práticas de *crowdsourcing* buscam preservar a autoria e a escala 1:1 dos documentos históricos, ao mesmo tempo que, devido ao alcance das mídias digitais, pretendem conservar grande quantidade de novos documentos^{XXIV}. Para isso, muitos deles se servem de programas de fácil acesso, como Omeka, que disponibiliza formulários eletrônicos nos quais os usuários podem fornecer sua contribuição – assim como seus metadados – preenchendo, ao mesmo tempo, a função de testemunha e arquivista.

Com maior ou menor adesão a esse conjunto de práticas e *software*, as iniciativas de arquivamento da memória da pandemia têm adotado o *crowdsourcing* como método preferido para formar seus acervos. Isso é perceptível em um dos primeiros projetos de arquivamento da memória da pandemia levado a cabo por historiadores, o *Coronarchiv*, sediado na Universidade de Hamburgo, na Alemanha. O

ARQUIVO, MEMÓRIA E BIG DATA: UMA PROPOSTA A PARTIR DA COVID-19

IAN KISIL MARINO
PEDRO TELLES DA SILVEIRA
THIAGO LIMA NICODEMO

projeto é formado por três historiadores daquela instituição e se distribui entre um site e contas no Facebook, no Twitter e no Instagram, para maior engajamento com o público. No site, os usuários podem compartilhar suas experiências e seus registros em diferentes formatos de mídia, através de um formulário articulado, que inclui um título, os arquivos, uma descrição e as informações sobre a data de criação e/ou captura dos registros dos quais será feito o upload.

A mesma abordagem, ainda que com menor aporte tecnológico, vem sendo feita no Brasil, através das propostas capitaneadas pelo Arquivo Público do Estado do Rio Grande do Sul – *Documentando a Experiência da Covid-19 no Rio Grande do Sul* –, e pelo Arquivo Geral do Município do Rio de Janeiro, que disponibiliza formulário elaborado no Google para que os usuários possam compartilhar sua experiência. Este caso se destaca pelo detalhamento das questões, que incluem desde dados socioeconômicos dos usuários até as percepções de quando souberam da nova doença, ainda restrita na China. Muitas dessas iniciativas também se caracterizam pela interface com a história oral. Isso é mais destacado no projeto *Histórias da Pandemia na Baixada Fluminense*, realizado pelo Programa de Pós-Graduação em Informática e Humanidades Digitais da Universidade Federal Rural do Rio de Janeiro. A iniciativa disponibiliza um número de WhatsApp para que os usuários da região, muitos deles de baixa renda, enviem vídeos de até dois minutos com os seus relatos sobre a experiência da pandemia e os impactos sobre as suas realidades. O projeto se destaca pela atenção às necessidades do grupo específico ao qual atende e aos constrangimentos tecnológicos aos quais está suscetível.

Esses projetos são iniciativas engenhosas, que trabalham na interface entre as tecnologias digitais e a realidade social mais ampla. Nesse sentido, eles se caracterizam pela tentativa de criação e manutenção do estímulo a setores do público que, fora da plataforma na qual contribuem agora, têm em comum o fato de terem sido impactados por um mesmo evento, ainda que não se conheçam entre si. Este aspecto condiz com o ensinamento de Roy Rosenzweig e Daniel J. Cohen^{XXV}, comentando a criação do *September 11th Digital Archive*, para quem os arquivos digitais mais bem sucedidos eram aqueles que tratavam de temas “quentes” e conseguiam articular uma base de usuários que compartilhavam uma mesma experiência off-line.

Não obstante os méritos do *crowdsourcing*, no entanto, é preciso reconhecer que essas iniciativas tendem a mobilizar um público que já está predisposto a compartilhar suas experiências no arquivo, de modo que os registros pouco têm de espontâneo. Da mesma forma, no caso da pandemia do novo coronavírus, tão importante quanto aquilo que os usuários pretendem conservar é aquilo que, com o passar do tempo, gostariam de esquecer. O que dizer das inúmeras *fake news*, rumores e boatos que têm sido tão importantes nos últimos meses e cuja preservação, assim como o conhecimento de onde surgiram e como se difundiram, poderiam garantir a responsabilização judicial de seus autores? Esse museu digital da infâmia, para aludir ao conto de Jorge Luís Borges, resgata a separação já feita por Marc Bloch entre os testemunhos “voluntários” e os “involuntários”. A memória da Covid-19, acreditamos, não pode se restringir somente aos primeiros.

Uma proposta de memória para a pandemia

ARQUIVO, MEMÓRIA E BIG DATA: UMA PROPOSTA A PARTIR DA COVID-19

IAN KISIL MARINO
PEDRO TELLES DA SILVEIRA
THIAGO LIMA NICODEMO

A abordagem do *crowdsourcing*, como vimos, tem sido amplamente adotada, no Brasil e internacionalmente, como forma de estimular um processo de arquivamento colaborativo e em larga escala. Entretanto, a nossa proposta se caracteriza por não se enquadrar exclusivamente nessa abordagem. O *crowdsourcing*, apesar de atraente devido a sua familiaridade à linguagem das redes sociais e ao amplo potencial de colaboração que propicia, também está sujeito a dinâmicas negativas. Há o risco de solaparem-se determinadas experiências minoritárias, ainda que relevantes, em nome de outras mais alinhadas com os parâmetros de aceitação digital ali estabelecidos^{XXVI}; do mesmo modo, como destaca Byung-Chul Han^{XXVII}, a colaboração digital gera resultados mais de caráter aditivo do que criativo, na medida em que a linguagem das redes sociais privilegia a sobreposição de experiências singulares, dificultando a criação de uma narrativa, que é o motivo que garante a relevância do lembrar e do esquecer no processo arquivístico.

Procura-se aqui, portanto, não apenas criar um arquivo digital da memória da pandemia, mas compreender, em teoria, o que é a memória e qual o lugar dos historiadores no mundo digital contemporâneo, além de viabilizar ferramentas e procedimentos para que sejam criadas as condições para o tratamento da memória como um patrimônio público da humanidade nesse cenário.

Para isso, almeja-se a criação de ferramentas digitais originais por meio de vínculos transdisciplinares inovadores, os quais possibilitarão a ressignificação do arquivamento e do trabalho dos historiadores face às carências e aos problemas contemporâneos. A partir da parceria com o projeto *DéjàVu*, do Instituto de Computação da Unicamp, dedicado à elaboração de ferramentas de computação forense em âmbito digital, busca-se analisar os resultados do projeto a partir do desenvolvimento de um aparato automatizado para o mapeamento da circulação das publicações, principalmente das *fake news*. Busca-se estabelecer as relações filogenéticas no interior de um banco de dados de publicações, a partir do cruzamento tanto dos metadados desses registros, como autoria e data de publicação, quanto da análise semântica do conteúdo das publicações. A partir de ferramenta de identificação própria das redes sociais e do tratamento de dados NoSQL (Non Structured Query Language, linguagem de busca não estruturada), pretende-se conseguir mapear a circulação de notícias e rumores sobre a pandemia, possibilitando saber quando e por que meios determinados conteúdos circularam, quais as reações que engendraram entre os usuários e quais os públicos mais afetados. Essa ferramenta possibilitará, como afirmamos acima, preservar também o que não deveria ser lembrado, mas nem por isso foi menos importante.

Junto a isso, busca-se criar um algoritmo que auxilie o trabalho dos historiadores – e dos arquivistas – ao lidarem com documentos digitais em larga escala. Novamente elaborada em parceria com o *DéjàVu*, será desenvolvido um algoritmo de detecção de relevância, capaz de mobilizar um subconjunto de itens úteis para análises quantitativas e qualitativas próprias à pesquisa historiográfica. Considerando a imensa quantidade de registros analisados, a ferramenta emularia o trabalho do historiador, principalmente na classificação de fontes, em uma escala que seria inviável em tão pouco tempo para qualquer indivíduo. O desenvolvimento dessa ferramenta demonstra o campo de colaborações entre a história e a ciência da computação, na medida em que seu aperfeiçoamento depende da análise da relevância feita pelo historiador, que a alimentará constantemente, treinando-a para reconhecer as relações e os contornos de

relevância mais importantes para o trabalho historiográfico. Conforme seu aprimoramento, o algoritmo terá o potencial de viabilizar uma leitura de fontes armazenadas em Big Data, engendrando possibilidades de pesquisa a partir de um fenômeno de escala global – a pandemia do novo coronavírus –, mas também semeando frutos para as pesquisas historiográficas futuras.

O que está em jogo é a questão incipiente, porém incontornável, do lugar da história e da memória no mundo contemporâneo. A pandemia da Covid-19 é um evento disruptivo global que, ao mesmo tempo, acelerou o processo de internalização das relações sociais na lógica digital e desnudou o problema das formas e do agenciamento sobre o lembrar e o esquecer nos arquivos digitais. A criação de um banco de dados com essas memórias, o estudo de experiências de arquivos digitais contemporâneos análogos, o desenvolvimento de aparatos digitais originais, a estruturação de um laboratório transdisciplinar especializado e a realização de parcerias com instituições públicas de pesquisa e de ensino de ponta caminham na direção da busca de uma epistemologia da história em nossa contemporaneidade digital.

História e o dever de memória em tempos de pandemia

Não é exagero afirmar que a pandemia é um evento transformador. Ninguém sabe quando ou se voltaremos ao normal, tampouco que “normal” será esse. Enquanto as universidades, na sua maioria, não retomam as atividades presenciais, mais e mais esferas da nossa vida cotidiana são absorvidas pelas tecnologias digitais. É um testemunho à imprevisibilidade da história que Peter Turchin, citado na abertura do texto, não tenha conseguido prever um evento de tamanha magnitude – embora, também, seja um pouco perturbador que ele tenha “acertado” que 2020 seria um ano de comoção e inquietação públicas acima do normal.

O jogo entre passado e futuro, porém, não se desenrola apenas no âmbito do processo histórico, mas também na definição da identidade do saber histórico, conforme ele varia com o tempo. Valores como a capacidade de reflexão e problematização, a adoção de uma perspectiva crítica e a busca por enunciados mais responsáveis, balizados com os princípios do método histórico, não precisam ser abandonados – tampouco defendidos de maneira conservadora – apenas porque a nossa paisagem tecnológica, midiática e disciplinar mudou. Pelo contrário, deve-se avaliá-los, julgá-los e atualizá-los conforme as demandas de cada época.

Eles também não precisam ser descartados porque as plataformas que possibilitam a comunicação digital pouco lhes dão espaço. As mídias digitais possuem, de certo modo, uma linguagem própria, baseada em protocolos numéricos e algorítmicos pré-formulados por profissionais da área da tecnologia de informação. Esses protocolos permitem a operação mecânica da projeção de dados, o que tem efeitos tanto positivos quanto negativos no campo da preservação patrimonial. É positivo que essas mídias proporcionem uma capacidade de colaboração, velocidade e acessibilidade sem igual na história da experiência arquivística; mas é negativo que, no geral, os fundamentos disciplinares da história e da arquivística não constem nesses protocolos digitais. Assim, preocupações como os critérios sobre a lembrança e o esquecimento, a classificação e a ordenação das informações e o agenciamento sobre as memórias criadas pela circulação social desses registros são delegados a profissionais cuja competência técnica não é acompanhada pela reflexão teórica, quando não são objeto de

ARQUIVO, MEMÓRIA E BIG DATA: UMA PROPOSTA A PARTIR DA COVID-19

IAN KISIL MARINO
PEDRO TELLES DA SILVEIRA
THIAGO LIMA NICODEMO

preocupações puramente mercadológicas, com todos os danos que a falta de moderação do conteúdo online tem causado às nossas democracias.

A experiência da Covid-19, com as suas milhares de memórias em profusão, não pode ser confiada cegamente às operações digitais. O digital deve ser o vetor da criação e divulgação de um acervo da memória da pandemia, mas o controle e os critérios desse arquivo devem garantir a confiabilidade e o interesse público de um legado que é dos mais significativos dos últimos tempos. Os historiadores têm sido os profissionais dedicados às representações das transformações da experiência humana no tempo, viabilizando ferramentas para sua preservação e constante ressignificação. É impensável, assim, que os historiadores não participem ativamente do cuidado do patrimônio legado pela pandemia.

Notas

^I Mestrando. UNICAMP. Pesquisador no Centro de Humanidades Digitais IFCH-UNICAMP. E-mail: ian.marino@usp.br

^{II} Pós-Doutorando. UNICAMP. Pesquisador no Centro de Humanidades Digitais IFCH UNICAMP. E-mail: pedrotellesdasilveira@gmail.com

^{III} Pós-Doutor e Professor de Teoria da História. UNICAMP. Pesquisador no Centro de Humanidades Digitais IFCH-UNICAMP. E-mail: tnicodem@unicamp.br

^{IV} SIMONDON, Gilbert. **On the Mode of Existence of Technical Objects**. Minneapolis: Univocal, 2017, p. 15.

^V GULDI, Jo; ARMITAGE, David. **The History Manifesto**. Cambridge: Cambridge University Press, 2017.

^{VI} COHEN, Daniel J.; ROSENZWEIG, Roy. "Collecting History Online", in ROSENZWEIG, Roy. **Clio Wired: The Future of the Past in the Digital Age**. New York: Columbia University Press, 2011, pp. 124-150.

^{VII} NOIRET, Serge. "História Pública Digital", in *Liinc em Revista*, Rio de Janeiro, v.11, n.1, maio 2015., pp. 28-51; FICKERS, Andreas. "Towards A New Digital Historicism? Doing History In The Age Of Abundance", in *Journal of European History and Culture*, Vol. 1, 1, 2012, pp. 19-26.

^{VIII} KOSELLECK, Reinhart et al. **O conceito de história**. Belo Horizonte: Autêntica, 2019.

^{IX} FOUCAULT, Michel. **A arqueologia do saber**. Rio de Janeiro: Relume Dumará, 1986; DERRIDA, Jacques. **Mal de arquivo: uma impressão freudiana**. Rio de Janeiro: Relume Dumará, 2001.

^X FICKERS, Andreas. "Towards A New Digital Historicism? Doing History In The Age Of Abundance", in *Journal of European History and Culture*, Vol. 1, 1, 2012, pp. 19-26, p. 4.

^{XI} WIMMER, Mario. "The Present as Future Past: Anonymous History of Historical Times", in *Storia della Storiografia*. n. 68. 2015, p. 77.

^{XII} STOLER, Ann Laura. "Colonial Archives and the Arts of Governance: On the Content in the Form", in HAMILTON, Carolyn et al. **Refiguring the Archive**. Dordrecht: Springer, 2002, p. 83-102.

^{XIII} FICKERS, Andreas. "Towards A New Digital Historicism? Doing History In The Age Of Abundance", in *Journal of European History and Culture*, Vol. 1, 1, 2012, p. 4; VAJCNER, Mark. "The Importance of Context for Digitized Archival Collections". *Ann Arbor: MPublishing*, University of Michigan Library. vol. 11, no. 1, April 2008, n. p.

^{XIV} BLOM, Ina. "Rethinking Social Memory: Archives, Technology, and the Social", in BLOM, Ina et al (eds). **Memory in Motion: Archives, Technology and the Social Memory**. Amsterdam University Press, 2017, p. 12, (tradução nossa).

^{XV} MOSS, Michael. "Opening Pandora's Box: What is an archive in the digital environment?", in CRAVEN, Loise. **What are Archives?** Burlington: Ashgate, 2008, pp. 77-88, (tradução nossa).

^{XVI} MOROZOV, Evgeny. **Big Tech: ascensão dos dados e a morte da política**. São Paulo: Ubu Editora, 2018.

^{XVII} BRÜGGER, Niels; SCHROEDER, Ralph. "Introduction: The Web as History", in BRÜGGER, Niels; SCHROEDER, Ralph (eds). *The Web as History. Using Web Archives to Understand the Past and the Present*. Londres: UCL Press, 2017, p. 10.

^{xviii} MOROZOV, Evgeny. **Big Tech**: ascensão dos dados e a morte da política. São Paulo: Ubu Editora, 2018, p.29.

^{xix} HESS, Aaron. “In digital remembrance: vernacular memory and the rhetorical construction of web memorials”, in *Media, Culture & Society*, vol. 29 (5), 2007, pp. 812-830; JARVIS, Lee. “Remember, remember 11 September: memorializing 9/11 on the Internet”, in *Journal of War & Culture Studies*, vol. 3, nº 1, 2010, pp. 69-82.

^{xx} KEATING, Joshua. “Covid-19 Is the First Truly Global Event”, in *The Slate*, May 5th, 2020, disponível em <https://slate.com/news-and-politics/2020/05/covid-19-global-event-pandemic.html>.

^{xxi} NORA, Pierre. “O retorno do fato”, in NORA, Pierre; LE GOFF, Jacques (orgs.). **Fazer história**: novos problemas. Amadora: Bertrand, 1977, pp. 179-193; FOGU, Claudio. “Digitalizing Historical Consciousness”, in *History & Theory*, vol. 48, issue 2, May 2009, pp. 103-121;

DOSSE, François. **Renascimento do acontecimento**. Um desafio para o historiador: entre Esfinge e Fênix. São Paulo: Editora da UNESP, 2013.

^{xxii} BRENNAN; T. Mills KELLY, Sheila E. . “Why Collecting History Online is Web 1.5”, in Roy Rosenzweig Center for History and New Media, 2009, disponível em: <http://chnm.gmu.edu/essays-on-history-new-media/essays/?essayid=47>. (tradução nossa).

^{xxiii} ERNST, Wolfgang. **Digital Memory and the Archive**. Minneapolis: University of Minnesota Press, 2013, p. 95.

^{xxiv} RIDGE, Mia (ed.). **Crowdsourcing Our Cultural Heritage**: Digital Research in the Arts and the Humanities. Farnham: Ashgate, 2014.

^{xxv} COHEN, Daniel J.; ROSENZWEIG, Roy. “Collecting History Online”, in ROSENZWEIG, Roy. **Clio Wired**: The Future of the Past in the Digital Age. New York: Columbia University Press, 2011, pp. 124-150.

^{xxvi} MOROZOV, Evgeny. **Big Tech**: ascensão dos dados e a morte da política. São Paulo: Ubu Editora, 2018, pp. 27-28.

^{xxvii} HAN, Byun-Chul. **No enxame**: perspectivas do digital. Petrópolis: Vozes, 2018, pp. 42-43.

Referências bibliográficas

BLOM, Ina. “Rethinking Social Memory: Archives, Technology, and the Social”, in BLOM, Ina et al (eds). **Memory in Motion**: Archives, Technology and the Social Memory. Amsterdam University Press, 2017, pp. 11-38.

BRENNAN; T. Mills KELLY, Sheila E. . “Why Collecting History Online is Web 1.5”, in Roy Rosenzweig Center for History and New Media, 2009, disponível em <http://chnm.gmu.edu/essays-on-history-new-media/essays/?essayid=47>.

BRÜGGER, Niels; SCHROEDER, Ralph. “Introduction: The Web as History”, in BRÜGGER, Niels; SCHROEDER, Ralph (eds). *The Web as History. Using Web Archives to Understand the Past and the Present*. Londres: UCL Press, 2017.

COHEN, Daniel J.; ROSENZWEIG, Roy. “Collecting History Online”, in ROSENZWEIG, Roy. **Clio Wired**: The Future of the Past in the Digital Age. New York: Columbia University Press, 2011, pp. 124-150.

DERRIDA, Jacques. **Mal de arquivo**: uma impressão freudiana. Rio de Janeiro: Relume Dumará, 2001.

DOSSE, François. **Renascimento do acontecimento**. Um desafio para o historiador: entre Esfinge e Fênix. São Paulo: Editora da UNESP, 2013.

ERNST, Wolfgang. **Digital Memory and the Archive**. Minneapolis: University of Minnesota Press, 2013.

FICKERS, Andreas. “Towards A New Digital Historicism? Doing History In The Age Of Abundance”, in *Journal of European History and Culture*, Vol. 1, 1, 2012, pp. 19-26.

FOGU, Claudio. “Digitalizing Historical Consciousness”, in *History & Theory*, vol. 48, issue 2, May 2009, pp. 103-121.

FOUCAULT, Michel. **A arqueologia do saber**. Rio de Janeiro: Relume Dumará, 1986.

GULDI, Jo; ARMITAGE, David. **The History Manifesto**. Cambridge: Cambridge University Press, 2017.

HAN, Byun-Chul. **No enxame: perspectivas do digital**. Petrópolis: Vozes, 2018.

HESS, Aaron. “In digital remembrance: vernacular memory and the rhetorical construction of web memorials”, in *Media, Culture & Society*, vol. 29 (5), 2007, pp. 812-830.

JARVIS, Lee. “Remember, remember 11 September: memorializing 9/11 on the Internet”, in *Journal of War & Culture Studies*, vol. 3, nº 1, 2010, pp. 69-82.

KEATING, Joshua. “Covid-19 Is the First Truly Global Event”, in *The Slate*, May 5th, 2020, disponível em <https://slate.com/news-and-politics/2020/05/covid-19-global-event-pandemic.html>.

KOSELLECK, Reinhart et al. **O conceito de história**. Belo Horizonte: Autêntica, 2019.

MOROZOV, Evgeny. **Big Tech: ascensão dos dados e a morte da política**. São Paulo: Ubu Editora, 2018.

MOROZOV, Evgeny. **To save everything click here**. Nova Iorque: PublicAffairs, 2013.

MOSS, Michael. “Opening Pandora’s Box: What is an archive in the digital environment?”, in CRAVEN, Loise. **What are Archives?** Burlington: Ashgate, 2008, pp. 71-89.

NOIRET, Serge. “História Pública Digital”, in *Liinc em Revista*, Rio de Janeiro, v.11, n.1, maio 2015, pp. 28-51.

NORA, Pierre. “O retorno do fato”, in NORA, Pierre; LE GOFF, Jacques (orgs.). **Fazer história**: novos problemas. Amadora: Bertrand, 1977, pp. 179-193.

RIDGE, Mia (ed.). **Crowdsourcing Our Cultural Heritage**: Digital Research in the Arts and the Humanities. Farnham: Ashgate, 2014.

BRENNAN; T. Mills KELLY, Sheila E. . “Why Collecting History Online is Web 1.5”, in Roy Rosenzweig Center for History and New Media, 2009, disponível em <http://chnm.gmu.edu/essays-on-history-new-media/essays/?essayid=47>.

SIMONDON, Gilbert. **On the Mode of Existence of Technical Objects**. Minneapolis: Univocal, 2017.

SPINNEY, Laura. “History as a giant data set: how analysing the past could help save the future”, in *The Guardian*, November 12th 2019, disponível em <https://www.theguardian.com/technology/2019/nov/12/history-as-a-giant-data-set-how-analysing-the-past-could-help-save-the-future>.

STOLER, Ann Laura. “Colonial Archives and the Arts of Governance: On the Content in the Form”, in HAMILTON, Carolyn et al. **Refiguring the Archive**. Dordrecht: Springer, 2002, p. 83-102.

VAJCNER, Mark. “The Importance of Context for Digitized Archival Collections”. *Ann Arbor: MPublishing*, University of Michigan Library. vol. 11, no. 1, April 2008, n. p.

WIMMER, Mario. “The Present as Future Past: Anonymous History of Historical Times”, in *Storia della Storiografia*. n. 68. 2015, pp. 165-18.