

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

LICENCIATURA EM HISTÓRIA

LEANDRO FERREIRA SOUZA

A ESCRITA DA HISTÓRIA E O DIGITAL:

O USO DE FERRAMENTAS DAS HUMANIDADES DIGITAIS E DE FONTES

NASCIDAS DIGITAIS A PARTIR DO *TWITTER*

PORTO ALEGRE

2022

LEANDRO FERREIRA SOUZA

A ESCRITA DA HISTÓRIA E O DIGITAL:
O USO DE FERRAMENTAS DAS HUMANIDADES DIGITAIS E DE FONTES
NASCIDAS DIGITAIS A PARTIR DO *TWITTER*

Monografia de conclusão apresentada no curso de Licenciatura em História, do Instituto de Filosofia e Ciências Humanas da Universidade Federal do Rio Grande do Sul como requisito para obtenção do título de Licenciado em História.

ORIENTADORA: PROF.^a DR.^a CAROLINE
SILVEIRA BAUER

PORTO ALEGRE

2022

CIP - Catalogação na Publicação

Souza, Leandro Ferreira

A escrita da História e o digital: o uso de ferramentas das humanidades digitais e de fontes nascidas digitais a partir do Twitter / Leandro Ferreira Souza. -- 2022.

59 f.

Orientadora: Caroline Silveira Bauer.

Trabalho de conclusão de curso (Graduação) -- Universidade Federal do Rio Grande do Sul, Instituto de Filosofia e Ciências Humanas, Licenciatura em História, Porto Alegre, BR-RS, 2022.

1. História digital. 2. Humanidades digitais. 3. Teoria da história. 4. Metodologia da história. 5. História pública. I. Bauer, Caroline Silveira, orient. II. Título.

AGRADECIMENTOS

Mario Quintana disse que, se não fosse o Van Gogh, o que seria do amarelo? E eu pergunto: se não fossem vocês, o que seria de mim?

Primeira e principalmente, aos meus familiares, que não há agradecimento maior a lhes fazer. Aos meus pais, Sandra e Mateus e meu irmão e minha cunhada, Evandro e Alessandra, em especial, o núcleo que me apoiou durante toda a graduação. Sem vocês, teria sido mais difícil. Sei da minha sorte e do meu privilégio de ter vocês por perto - e de saber que sempre estarão.

Aos/às amigos/as que muito ouviram eu falando que tinha terminado mais uma página da monografia e muitas vezes me ajudaram a dispersar desses momentos de escrita de outras formas, obrigado a todos/as. Vocês são parte vital de quem eu sou e quem eu me tornei durante os últimos anos - e, alguns, desde criança.

Aos/às amigos/as e colegas da universidade, principalmente Vander, Cris, Nati e Elisa, que além do curso da graduação, muitos/as foram companheiros/as nas frustrações e satisfações deste momento de escrita da monografia. Em especial, deixo um agradecimento à Nati, pessoa que mais ouviu minhas reclamações durante esse processo, obrigado pela paciência de ouvir os infinitos áudios; obrigado por, por vezes, também ser minha corretora e dar opiniões sinceras sobre trechos da escrita; e obrigado por estar sempre por perto, tua companhia foi essencial. E à Elisa, minha eterna dupla da graduação (e de concurso [risos]), que me acompanhou (e, por vezes, salvou) em quase todos os trabalhos e disciplinas, compartilhando opiniões, gostos e desgostos nesse caminho, obrigado por me ouvir reclamar das coisas que nem precisavam de reclamação e compartilhar tuas próprias alegrias e insatisfações; que estejamos sempre caminhando juntos, tendo trocas como essas, e espero que saiba que a tua participação foi fundamental em tudo isso, até chegarmos aqui. Tu é absolutamente brilhante e o trabalho de conclusão de curso é só mais uma etapa da vida que tenho o prazer de dividir contigo, conta sempre comigo.

À professora Caroline Silveira Bauer cabe um agradecimento de, pelo menos, quatro anos. É sorte e honra minhas ter entrado no curso de história e, desde o segundo semestre, poder ser orientado por ti. É mais sorte ainda saber que além da relação de orientação, posso te considerar uma amiga. Nada disso, inclusive minha virada à história pública e digital e à linha de pesquisa que desejo seguir daqui pra frente, seria possível sem a tua influência e do LUPPA. Tua presença e acompanhamento foram essenciais pra toda minha formação: das bolsas de iniciação científica e popularização da ciência às manhãs e tardes assistindo e comentando virtualmente a CPI da Covid, tu ultrapassa - e muito - a simples condição de

orientadora. Que a tua tranquilidade, influência, dedicação e amor pela história alcancem o mundo. Enquanto pessoa e professora, te admiro profundamente.

Decerto, há muitas pessoas a quem eu agradeceria nominalmente aqui, mas careço de espaço para fazer os agradecimentos devidos a cada um de vocês - e eu não sou muito acostumado em resumir sentimentos em texto. Vocês são, todos e todas, incríveis. Finalmente, a todos/as que estão próximos à mim, amo vocês e deixo a brega mensagem da música de Dionne Warwick: *keep smiling, keep shining / knowing you can always count on me, for sure / that's what friends are for.*

RESUMO

Esta monografia propõe a reflexão sobre mudanças na escrita e teoria da história a partir de fontes nascidas digitais com o uso de ferramentas digitais-computacionais vinculadas às humanidades digitais. O objetivo é apresentar debates sobre o andamento dos campos de história digital e humanidades digitais, explicitando seus compromissos e atuações para, então, apresentar as ferramentas que fazem parte desse escopo e demonstrar seu uso na área das ciências humanas, com a finalidade de produzir uma forma de visualização em grafo. Utiliza-se a linguagem de programação *Python*, bem como o *software* Gephi para produzir uma raspagem e tratamento de dados oriundos da rede social virtual *Twitter*. Argumenta-se que esses instrumentos alteram a forma de escrever história, uma vez que possibilitam a criação de fontes nascidas no ambiente digital que servirão de fontes primária e secundária a ponto de modificar a metodologia da história, sem alterar o método científico da disciplina. Dessa forma, é possível compreender que os grafos, de certa forma, possibilitam diferentes interpretações sobre os dados brutos e, a priori, são catalisadores na criação de fontes primárias oriundas da internet e do ciberespaço, potencializando a escrita da história na era digital.

Palavras-chave: história digital; humanidades digitais; teoria da história; escrita da história; grafo; programação; ciências da computação; visualização de dados; fontes nascidas digitais; internet.

ABSTRACT

This paper proposes a reflection about the changes in the writing of history and historiography from born-digital archives with the usage of digital tools bound to the digital humanities. The aim is to present debates on the course of the digital history and digital humanities fields, explaining their commitments and procedures in order to introduce the tools that make part in this sphere and demonstrate its usage in the human sciences field, aiming to produce a graph form of data visualization. We use the Python programming language as well as the data analysis software Gephi to produce data mining and analysis of datasets extracted from social media Twitter. We argue that these instruments change the way of writing history once it makes the creation of born-digital archives on the digital environment possible, and that it will serve as primary and/or secondary sources and modify the historiography methodology without changing the discipline's scientific methods. Thus, it is possible to comprehend that graphs, in a certain way, possibilitate different interpretations about the raw data and could be catalyzers on the nurture of primary sources born on the internet and cyberspaces, improving the writing of history in the digital age.

Keywords: digital history; digital humanities; historiography; writing of history; graph; programming; computer science; data visualization, born-digital archive; internet.

SUMÁRIO

1. INTRODUÇÃO	9
2. A PRÁTICA HISTÓRICA EM TRANSFORMAÇÃO: as humanidades digitais, a historiografia e a internet	15
3. “MULTIDISCIPLINARIDADE DIGITAL”: as linguagens de programação e a ciência computacional agregadas às humanidades digitais.	26
4. DADOS DA WEB SEMÂNTICA, SUA VISUALIZAÇÃO ATRAVÉS DE GRAFOS E AS FONTES NASCIDAS DIGITAIS.	35
4.1. Como produzir grafos	40
4.2. Usos dos grafos na historiografia	45
5. CONSIDERAÇÕES FINAIS	50
6. REFERÊNCIAS BIBLIOGRÁFICAS	55

1. INTRODUÇÃO

Finalmente, o que é uma “obra de valor” em história? Aquela que é reconhecida como tal pelos pares. Aquela que pode ser situada num conjunto operativo. Aquela que representa um progresso com relação ao estatuto atual dos “objetos” e dos métodos históricos e, que, ligada ao meio no qual se elabora, torna possíveis, por sua vez, novas pesquisas. (DE CERTEAU, 2013, p. 64)

O historiador Pedro Telles da Silveira, na introdução de sua tese, ao falar sobre o *fazer história* no campo da história digital (HD), é direto ao afirmar que "tudo mudou, ainda que nada essencialmente tenha se alterado" (2018, p. 15). Essa curta frase é muito poderosa e significativa quando fala-se desse campo, tão recente em vista de toda a história e ainda bastante desconhecido por muitos e muitas historiadores/as. Na sequência, ele ainda escreve que "as transformações mais visíveis operadas pela introdução das novas tecnologias sobre o conhecimento histórico se manifestaram nesses atos diminutos que compõem os aspectos pragmáticos – no caso, prosaicos – da atuação dos pesquisadores e pesquisadoras" (2018, p. 16). O que isso significa, na práxis do/a historiador/a, é a intocabilidade no que compõe a o método científico da historiografia, muito embora este ainda sofra alterações para particularizar a forma de pesquisar e escrever história na era digital. Fabio Chang de Almeida (2011, p. 25) corrobora essa ideia ao falar que a história digital precisa de uma metodologia particular já calcada nas bases metodológicas da pesquisa historiográfica existente. Entretanto, é imprescindível citar aqui que, embora a metodologia não seja radicalmente alterada, suas mudanças acontecem de forma a lançar luz a novas formas de pesquisa, novos conceitos e novos entendimentos sobre o que é e como é **fazer história**.

A história digital enquanto um campo teórico-metodológico, dessa forma, carrega algumas importantes transições e novidades quando comparada à operação historiográfica anterior. Sem dúvidas, dessas novidades, a mais engajada nos debates sobre a HD é a digitalização de arquivos, museus e documentos históricos existentes no ambiente físico, tornando assim o acesso a esses documentos muito mais rápido e fácil. O olhar no horizonte, porém, nos permite perceber que, para além disso, há outras importantes contribuições provindas da HD. Para entender isso, é preciso a compreensão da história digital como um território anexo a um outro, maior: as humanidades digitais. Seu surgimento tem origem na junção das ciências humanas ou linguísticas com as ciências informacionais e computacionais, ou seja, é a compreensão da transdisciplinaridade de pelo menos duas áreas diferentes que constrói um campo de estudo imerso na tecnologia da informação e comunicação. É, portanto, a intersecção simultânea de instrumentos de diferentes áreas com

as perspectivas do universo digital. Nestes aspectos, as humanidades e a história digital têm constante intersecção entre as "nossas" práticas com aquelas das áreas da computação e da tecnologia da informação - bem como temos compartilhado termos e ideias desses campos. Percebe-se, assim, que, com relação às novidades supracitadas, há esse entrelace com novas áreas e a aproximação do 'nosso' método científico com o 'deles', bem como esse compartilhamento de tecnologias das respectivas áreas que passam a colaborar entre si. Entre eles, e para a produção deste trabalho, utilizamos, de forma contínua, termos como *big data*, *data analysis*, *data mining*, *data set*, bem como com *linguagem e códigos de programação*. Assim, entendemos que há a necessidade de algumas explicações prévias.

Com o advento da computação, dados, nas suas mais diferentes configurações, passaram a ser coletados, armazenados e analisados de forma diferente e em diferente volume (se comparado aos períodos anteriores). E, a cada evolução tecnológica que temos, descobrimos novas e mais fáceis formas de fazer este trabalho - por exemplo: desde 1987 existe o *software* da *Microsoft* conhecido como *Excel*, um serviço de planilhas operado em computadores que facilita o trabalho de milhões de pessoas em diferentes áreas, servindo como produto matemático, organizacional, de gerenciamento de dados, etc. À medida que essas evoluções ocorrem no mundo digital, os dados também "evoluem", portanto, hoje pode-se observar que a produção de dados na internet é colossal, de forma a não haver uma máquina capaz de processar todos os dados da internet de uma vez só.

O *Big Data* é, simplificada, esse conjunto de dados (*dataset*), maior e mais complexo, sendo coletado e armazenado constantemente, em uma enorme velocidade - tudo isso, claro, na rede mundial de internet. Empresas de tecnologia utilizam do *Big Data* como forma de analisar constantemente os dados de usuários, clientes potenciais, e produtos, de forma a precisar suas ações de negócios. Para esta pesquisa, o *Big Data* serve como o exemplo de facilidades que a internet pode trazer para a historiografia - nesse ínterim, podemos citar sobre como a introdução de computadores na pesquisa histórica no Brasil, em 1970 e 1980, possibilitou o trabalho com séries, que demorariam meses ou até anos para serem analisadas, se feitas manualmente; hoje, passados cinquenta anos, isso é feito em segundos. E é a partir desses dados que entram as noções de *data mining* e *data analysis*. Para um/a historiador/a comum, que não possui uma grande empresa de tecnologia e uma equipe de profissionais das áreas computacionais/informacionais ao seu dispor, não é tão fácil coletar esses imensos dados a todo momento. Logo, o que o conceito de *data mining* carrega é a possibilidade de fazermos, de forma automatizada, essa coleta - *data mining*, pode ser

chamado de “raspagem de dados”. A programação computacional é a responsável por essa automatização na coleta de dados por um usuário. E, com esses dados "em mãos", o/a pesquisador/a pode (e deve) analisá-los (chamado de *data analysis*), de diversas maneiras diferentes; a maneira de análise de dados que este trabalho pretende mostrar é a que acontece com grafos, uma forma de visualização de dados.

Dessa forma, um dos pontos principais que esta pesquisa pretende abordar é essa multidisciplinaridade do campo de pesquisa, alguns conceitos mútuos entre as diferentes áreas e a existência das fontes nascidas digitais (as fontes que são geradas digitalmente) e o uso de ferramentas computacionais das humanidades digitais. Neste íterim, quando falamos de fontes nascidas digitais, há especial atenção àquelas geradas a partir de redes sociais virtuais. Assim, a temática a ser explorada pelo projeto em questão envolve a reflexão sobre o uso das ferramentas computacionais das humanidades digitais para a escrita da história, com a utilização de fontes nascidas digitais. Essa caracterização se dá a partir da produção e uso de dados oriundos destas, tendo como pedestal metodológico a rede social *Twitter* e os dados que seu funcionamento produz, como publicações, compartilhamentos, curtidas, *hashtags*, *trending topics*, etc. As indispensáveis epistemologias que congregam esses diferentes saberes científicos contidos na multidisciplinaridade das humanidades digitais e que compreendem essa pesquisa passam, portanto, por algumas noções supracitadas, a ver: as linguagens de programação computacional para raspagem de dados; as ferramentas de análise e interpretação de dados; as redes sociais virtuais; as fontes nascidas digitais; e a cultura digital e histórica.

Para a produção de uma reflexão acerca dos usos epistemológicos da história digital, a utilização de um trabalho que contenha uma “ruptura” com o que pretendemos desenvolver com a digitalização computacional da historiografia é importante. Para isso, optamos por utilizar o segundo capítulo do livro do historiador Michel de Certeau, intitulado “A Operação Historiográfica”, presente na obra “A Escrita da História”. Sendo esta uma obra cuja primeira edição no Brasil foi publicada em 1982, entendo-a como pensada integralmente a partir da ótica “tradicional”¹ da historiografia, jamais para um campo de pesquisa como o digital. O

¹ Em alguns momentos deste trabalho se fará menção a uma forma “tradicional” da história. Para evitar interpretações divergentes, permito-me esclarecer o que entendo por história ou historiadores/as “tradicional”, quando comparados/as aos/as digitais. Compreendo o tradicional aqui enquanto o método histórico que não se utiliza de muito mais do que o “básico” que os computadores e a internet oferecem, i.e., os editores de texto, planilhas e documentos/livros/artigos em .pdf; aquele método que não se vale, portanto, das fontes nascidas digitais, tampouco das ferramentas computacionais ou outros *softwares* mais avançados, que requerem um usuário também avançado. Na falta de um melhor termo para caracterizar essa forma mais corriqueira de "escrita

autor inicia o capítulo com uma afirmação basilar: “a operação histórica se refere à combinação de um *lugar* social, de *práticas* “científicas” e de uma *escrita*” (DE CERTEAU, 2013, p.56). Atenho-me a aprofundar um raciocínio sobre as práticas, principalmente com a afirmação que “em história, tudo começa com o gesto de *separar*, de reunir, de transformar em “documentos” certos objetos distribuídos de outra maneira” (p. 73), justamente por esta ser uma afirmação adequada para um debate sobre as fontes nascidas digitais. Estas podem ser categoricamente chamadas de objetos distribuídos de outra maneira que, com o método e o cientificismo históricos, transformam-se em “documentos”. A operação historiográfica e o historiador *per se*, de acordo com o autor

[...] Trabalha sobre um material para transformá-lo em história. [...] Transformando inicialmente **matérias-primas** (uma informação primária) em produtos *standard* (informações secundárias), **ele os transporta de uma região da cultura** (as “curiosidades”, os arquivos, as coleções, etc.) **para outra (a história)**. (DE CERTEAU, 2013, p. 72, grifos meus em negrito)

A afirmação de De Certeau, além de extremamente didática, é perfeitamente adaptável ao fazer história do campo historiográfico digital, às ferramentas das humanidades digitais e às fontes nascidas digitais. No caso desta pesquisa, pretende-se a transformação de tais fontes (as matérias-primas) em informações e dados quantitativos e qualitativos que podem ser interpretados como os produtos *standards*, partindo de uma cultura digital histórica (as redes sociais virtuais) para, enfim, a história. A práxis é singelamente diferente, mas a episteme e a metodologia são, para todos os fins, tão cientificamente semelhantes quanto poderiam ser. Uma introdução do que contempla o fazer histórico ou historiográfico e essa operação que os/as historiadores/as produzem parte, assim, da ideia de debater sobre a metodologia

A seguir, para trabalhar com a noção de humanidades digitais propriamente, utilizaremos os/as autores/as Marco Antônio de Almeida e Ieda Pelógica Martins Damian, no seu trabalho "Humanidades Digitais: Um campo praxiológico para mediações e políticas culturais?" (2015), bem como Pedro Telles da Silveira e Bruno Grigoletti Laitano, em trabalhos mais recentes, intitulados, respectivamente: "História, técnica e novas mídias: reflexões sobre a história na era digital" (2018) e "Digitalizar o arquivo, arquivar o digital: a história e suas fontes diante das velhas e novas tecnologias" (2021). Estes são importantes trabalhos que, para além das discussões sobre as perspectivas do campo das humanidades digitais, também confluem uma série de outros trabalhos de importantes autores e autoras da área, como Anita Lucchesi, Matthew Kirschenbaum, Fabio Chang de Almeida, entre

da história", utilizei o termo "tradicional", embora ciente do peso que o conceito de tradição possua para a história.

outros/as. Aqui, tanto Telles quanto Laitano escrevem sobre a ideia de que as humanidades digitais não são simples - não é um campo que pode ser resumido ao uso de ferramentas digitais nas humanidades, ou apenas historiadores/as usando computadores e internet, ou então a ideia de "historiadores programadores"; é, talvez, a junção de tudo isso e mais algumas características, que envolvem prática e metodologia, e uma grande mobilização das tecnologias de informação no fazer histórico. Esse debate é justamente o que compõe parte dos trabalhos de Telles e Laitano, que, com esses diversos autores, mostram uma extensão teórica e pragmática em torno do próprio entendimento do que exatamente são as humanidades digitais e a história digital, diga-se.

Já o trabalho de Antônio de Almeida e Ieda Damian mostra que as "humanidades digitais" constroem um campo interdisciplinar servente às novas reflexões surgidas com a introdução de tecnologias digitais; defendem que, provavelmente, hoje há muitos profissionais, pesquisadores/as humanistas digitais que apenas não se rotulam dessa forma, mas a práxis de seus trabalhos o são. Essa é uma reflexão muito importante quando pensamos na resistência à digitalização da história (e das humanidades, certamente). Não é apenas porque a nossa área é, como defende Andreas Fickers (2012), conservadora, mas também porque o entendimento sobre o que faz a digitalização das humanidades ainda ser muito vago.

Com relação às ferramentas computacionais presentes nas humanidades digitais, a citar: o *Python* e o *R* - linguagens de programação - e o *Gephi* - um *software* de produção de grafos a partir de dados brutos -, são práticas difundidas e utilizadas por muitos/as pesquisadores/as e laboratórios de pesquisa, como o Laboratório de estudos sobre Imagem e Cibercultura (LABIC)², do Departamento de Comunicação Social da Universidade Federal do Espírito Santo, que produz grafos sobre diversos assuntos relevantes, envolvendo política, história, sociologia, estatísticas, etc., bem como o historiador e humanista digital francês *Martin Grandjean*, que utiliza grafos para inúmeros fins em seu blog³. Outro exemplo importante de um grande projeto envolvendo grafos é o *Mapping the Republic of Letters*⁴, produzido em *Stanford*, a fim de demonstrar em grafos as redes de comunicação da República das Letras, mapeando redes de correspondências entre letrados dos séculos XVII e XVIII.

As linguagens de programação citadas não são exclusivamente para a coleta de *um* tipo de dado, podendo ser adequadas para diversos trabalhos; há, no Brasil e no mundo,

² cf. Labic. Disponível em: <<http://www.labic.net/>>. Acesso em: 23 de novembro de 2021.

³ cf. Martin Grandjean: Digital humanities, Data visualization, Network analysis. Disponível em: <<http://www.martingrandjean.ch/?s=graph&search=Procurar>>. Acesso em: 23 de novembro de 2021

⁴ cf. Mapping the Republic of Letters. Disponível em: <<http://republicofletters.stanford.edu/>>. Acesso em: 23 de novembro de 2021.

historiadores/as utilizando a programação para produção de projetos de cartografia, de grafos, de produção de banco de dados ou simplesmente de coletas de informações a serem utilizadas de forma bruta (i.e., não tratada a partir de outra ótica) nas suas pesquisas. Os códigos de programação, neste caso, são o "meio" para se conseguir um produto final (dados).

Sobre o trabalho com dados, o autor Fabio Chang de Almeida (2011), bem como reforçado por Bruno Leal Pastor de Carvalho (2016) e Anita Lucchesi (2014), nos ajuda a compreender as possibilidades de utilização das redes sociais virtuais como produtoras de dados que podem despertar muitas outras problemáticas para a escrita da história, como a circulação da história no espaço público, os sentidos de história presentes nas redes sociais, os impactos de determinadas comemorações ou menções de fatos ou personagens históricos nas redes, etc.

Metodologicamente, a pesquisa partirá da revisão bibliográfica, buscando elencar as contribuições dos/as autores/as aqui citados e outros/as mais, que vêm trabalhando com esses temas, procurando sempre produzir uma reflexão alinhada aos debates mais recentes em torno das áreas e campos contemplados pela pesquisa. Depois de debatidas essas ideias, serão apresentadas as ferramentas e apontadas algumas de suas contribuições para a escrita da história. Por fim, é importante assinalar que, a fim de exemplificar sua potencialidade, serão feitos breves exercícios de raspagem de dados e de análise, com a utilização dessas ferramentas de humanidades digitais, de forma a não comprometer a temática apresentada no projeto de pesquisa, mas poder pragmaticamente mostrar sua práxis.

Essa pesquisa pretende ser pensada, interpretada e realizada a partir de duas dimensões: a relevância acadêmica e a relevância social presentes nela. Aquela, estritamente ligada à teoria, pretende compreender a potencialidade de inovação que a ideia pode trazer para a escrita da história; que tipo de análise pode ser produzida que outras fontes não possibilitam; como a automatização de tarefas pode também ser uma forma de inovação do trabalho tradicional do/a pesquisador/a em muitas vezes ter que, manualmente, recolher e organizar um grande número de informações; bem como a interdisciplinaridade, indissociável da digitalização da história⁵, uma vez que a utilização de áreas de atuação pertencentes à Ciência da Computação é imprescindível para obter melhores resultados. Já a relevância social carrega a participação intrínseca da história pública na forma de se relacionar com um

⁵ Julgamos necessário uma explicação prévia aqui sobre o que entendemos por “digitalização da História” não ser unicamente a digitalização dos papéis e dos documentos físicos como um todo, mas, sim, como ficará explícito ao longo deste projeto, a imersão da pesquisa histórica no mundo digital, conjuntamente com as ferramentas computacionais, as linguagens de programação, os *softwares* de uso avançado, enfim, toda a potencialidade que essa multidisciplinaridade traz.

público leigo e/ou não-acadêmico - dessa forma, a aproximação entre o/a historiador/a e seu público; essa aproximação está disposta, por exemplo, na construção de uma forma coletiva e colaborativa de produção da história, o que significa a participação pública e o saber comunitário ajustados ao trabalho do/a historiador/a, conceito este definidor da "história pública digital" para Serge Noiret (2015, p. 35); além disso, a necessidade de ampliar a compreensão da cultura digital-histórica e do desenvolvimento tecnológico que permite a emergência de um novo modelo que busca sair do analógico e adentrar um novo universo (digital) também é força-motriz dessa expressão social do projeto.

Dessa forma, esse trabalho pretende refletir sobre as possibilidades da combinação da práxis metodológica das humanidades digitais com as ferramentas digitais das ciências computacionais e informacionais para as pesquisas historiográficas, possibilitando a formulação de novas hipóteses e problemáticas, a partir do uso de duas dessas ferramentas e da exploração de fontes nascidas digitais produzidas no âmbito do *microblog Twitter*. As pretensões particulares são, portanto, i) compreender a potencialidade das ferramentas computacionais para a historiografia acadêmica; ii) explorar a possibilidade de aplicar um método de identificação e tratamento das fontes nascidas digitais enquanto fontes primárias; iii) ressaltar o uso dessas ferramentas como um dispositivo relevante para novos/as historiadores/as.

Divide-se, para isso, o presente trabalho em três partes: na primeira, buscaremos uma abordagem sobre a história digital e como as ferramentas computacionais e as fontes nascidas na internet mudaram a forma de escrever a historiografia; na segunda, pretendemos aproximar o/a cientista humano/a da área da computação, mostrando como a multidisciplinaridade pode ser um importante intercâmbio de informações e práxis entre nossas áreas; por fim, a terceira parte terá o papel de complementar o que as duas primeiras estão aprofundando, utilizando para isso exemplos práticos de uso de grafos e a produção de fontes oriundas do digital para a operação histórica.

2. A PRÁTICA HISTÓRICA EM TRANSFORMAÇÃO: as humanidades digitais, a historiografia e a internet

O relacionamento entre a história e as tecnologias digitais reproduz o ciclo de alternância entre otimismo e pessimismo que caracteriza a trajetória do uso da internet, e se relaciona com as tecnologias não apenas através do que elas podem fazer mas, mais significativamente, através do que se acredita que elas são capazes de realizar. (SILVEIRA, 2018, p. 30)

O constante avanço da tecnologia e as mudanças que invariavelmente vão alterando as nossas percepções sobre as mídias digitais trouxe à história chamada digital (e a seus/suas profissionais) uma dificuldade (que é, também, uma necessidade agradável): precisamos estar, sempre, atualizados/as sobre o que nossos/as pares estão fazendo⁶. Ou isso, ou não teremos controle sobre nosso próprio trabalho. Além disso - e principalmente -, na HD essa atualização do/a pesquisador/a é referente **também** às técnicas (conhecer e compreender novos *softwares*, *hardwares*, linguagens de programação e os seus usos dentro do campo da historiografia).

Em 2005, Daniel J. Cohen e Roy Rosenzweig⁷ se debruçaram sobre a hipótese da revolução digital *cyber*-histórica, em que teríamos uma nova, mais rica e mais profunda Era historiográfica. Nesse mesmo trabalho, os autores levantam diversas realidades advindas da digitalização da história até hoje debatidas - tanto é que fazemos questão de citar o trabalho deles. Cohen e Rosenzweig são, para todos os fins, grandes nomes que tornaram públicos os pensamentos sobre a necessidade de uma história digital e quais eram as suas dificuldades e facilidades, seus desafios e maneiras de lidar com essa nova forma de preservar o passado - pelo menos em 2005.

Entretanto, como fica claro perante a leitura do seu trabalho, em 2022 - ano em que redigimos este trabalho - não podemos nos ater somente às discussões estabelecidas pelos autores naquele ano. Estamos em um novo momento, com novas tecnologias e novos aparatos informacionais para nos auxiliar. Enquanto Cohen e Rosenzweig estabeleceram uma conexão entre a historiografia e o *XML*⁸, hoje já podemos fazer esta mesma conexão entre a historiografia e as linguagens de programação, os *softwares* de uso avançado, as inteligências artificiais ou *machine learnings*, i.e., novidades que, em 2005, sequer eram probabilidades para o nosso campo. Entretanto, isso não significa que todos/as os/as historiadores/as tenham dominado o XML ou a construção de bancos de dados através dessas ferramentas pertinentes, mas sim que o seu uso específico não é mais necessário uma vez que, globalmente, isso foi suplantado por outra ferramenta. Também cabe dizer que ainda há muito o que aprender, do

⁶ Certamente, estar atualizado/a deve ser prioridade de todo/a historiador/a pesquisador/a, mas há a relativa diferença entre alguns espaços já consolidados, onde as correntes teóricas mais atuais estão distribuídas em trabalhos de longa data. Dessa forma, estar inserido num campo de história do trabalho na modernidade, por exemplo, é diferente de estudar o tempo presente.

⁷ COHEN, D. J.; ROSENZWEIG, R. **Digital history: a guide to gathering, preserving, and presenting the past on the Web**. [S. l.]: University of Pennsylvania Press, [s. d.].

⁸ Sigla para *eXtensible Markup Language*, um modelo/parâmetro utilizado para codificar documentos (hoje muito utilizada para criação de Notas Fiscais Eletrônicas, as NFe) e criar banco de dados (ofício 'superado' pelo formato JSON, *JavaScript Object Notation*, hoje muito mais universalizado nas empresas)

ponto de vista dos/as historiadores/as que trabalham com bancos de dados e repositórios, o que deixa explícito como o debate sobre a construção de bancos de dados é atual, apenas a forma com a qual fazê-lo que não. Se trocarmos o XML por uma combinação atualizada de instrumentos que tenham a mesma finalidade, é uma discussão adequada.

Pensando no estado em que se encontra a arte e a filosofia da ciência da computação, da ciência da informação e das humanidades digitais, o campo historiográfico deve ser continuamente atualizado em suas possibilidades. O contínuo aperfeiçoamento tecnológico faz com que os/as historiadores/as do meio digital-informacional tenham que se manter atentos/as às inovações provenientes de diferentes disciplinas. Por isso, há diversos/as intelectuais debatendo e divulgando o conhecimento histórico ao passo que este estabelece novas relações com o corpus documental e com as novidades digitais-computacionais - tal como podem ser percebidas a digitalização da história e as fontes nascidas digitais. Daí, também, a manutenção de uma antiga provocação: o/a historiador/a do futuro será um/a programador/a, ou não será. Evidentemente, à parte dessa provocação irônica, a história e o/a historiador/a que é digital não se propõe simplesmente à escrita dos seus textos em um computador, tampouco à prática investigativa de documentos digitalizados nos arquivos e bibliotecas; é pertinente entender que a continuidade deste trabalho pressupõe inequivocamente a digitalização da história ser, a priori, a transformação do seu *modus operandi* a partir de ferramentas, *softwares* e dispositivos computacionais e a incorporação destes à prática própria do campo.

Alguns anos depois de Cohen e Rosenzweig, a historiadora público-digital Anita Lucchesi argumenta sobre como a cultura é diretamente influenciada pelo advento das tecnologias (2014, p. 46) e, com a cultura, a escrita da história também; a produção do conhecimento histórico de hoje possui ferramentas e dispositivos muito diferentes da década de 1970 ou mesmo do início do século XXI.

A autora também dialoga sobre como a transição de uma cultura analógica para uma digital faz surgir novas formas de escrever e ler, com novas linguagens (visíveis e invisíveis, verbais, orais e icônicas), para isso, exemplifica como um ícone de uma lixeira ou de um envelope, no Brasil ou no Japão, possuem o mesmo significado (2014, p. 47). Lucchesi demonstra, com essa ideia, o surgimento de novas linguagens universais, onde, por exemplo, o ícone de um envelope em um computador pertencente a um cidadão brasileiro é entendido da mesma maneira por um cidadão japonês (é o ícone relativo ao *email*); a mesma coisa

acontece com o ícone de lixeira (relativo ao local para onde arrastamos os arquivos que não queremos mais no nosso dispositivo). Emojis, de certa forma, funcionam da mesma forma, uma vez que um emoji sorrindo surte uma interpretação semelhante em diferentes países e culturas. Isso acontece também com as linguagens de programação, em que uma linha de código lida por um dispositivo no Brasil, na Austrália ou na Dinamarca vai ter o mesmo efeito, com a única ressalva de que esses códigos utilizam expressões em inglês (que, embora não sejam nossa língua materna, são o motivo da universalidade das linguagens de programação). Essa virada tecnológica e digital que acontece no mundo com o surgimento da *World Wide Web* é imprescindível para entender a transformação da historiografia de ontem e de hoje. Lucchesi, em um trabalho anterior, datado em 2012, também diz que, sobre a escrita da história, ela não deve ser simplesmente vista como a digitalização da história para as telas dos computadores e celulares, mas que:

Estamos falando da transformação dos meios, das novas mídias digitais, cujas qualidades que interessam destacar nos limites deste texto podem ser melhor compreendidas se pensadas nas suas dimensões hipertextuais, encerrada para alguns autores no conceito de hiperímia (LUCCHESI, 2012, p. 3)

E, defendendo a chamada hipertextualidade da escrita, i.e., a “possibilidade da escrita em camadas, da construção estratificada dos textos em vários níveis para além da aparente plana e linear superfície do nosso écran” (2012, p. 9), a historiadora expõe como essa característica pode transformar um texto em uma interação imaginada pelo próprio autor. É, também, uma construção de história pública com o público-leitor.

Surge a possibilidade da prática histórica alinhada aos recursos computacionais, midiáticos e visuais; surge, portanto, essa historiografia com tangibilidade no espaço digital, não determinada pela representação tradicional da história que se faz no arquivo, no museu ou na biblioteca. É desafiador, como há vinte anos, pensar nessa fuga a um cenário fisicamente intangível para um grupo de profissionais acostumados/as a perceber o mundo material como sua maior fonte de pesquisa. Para Lucchesi, ainda, a historiografia digital é “inscrita no ciberespaço, escrita digitalmente (hipertextualmente) e é divulgada na rede” (2014, p. 50). Junto a essas características, há ainda outras singularidades originadas nessa digitalização do campo, a citar: a empregabilidade das ferramentas computacionais ao fazer histórico, a aproximação entre a história e campos afastados das ciências sociais e humanas - como as ciências da comunicação e as ciências da computação - e a composição de novas metodologias científicas que só se fazem possíveis com e a partir a *world wide web*. Além disso, há a distinção mais direta, a abundância de fontes. Esta é citada por Andreas Fickers e

reforçada por Pedro Telles da Silveira, que cita que isso acabou “alterando a dinâmica da pesquisa, acelerando-a graças à sua maior facilidade de acesso” (2018, p. 49).

Sequencialmente, ao passo que Fickers argumenta sobre a história ser uma área ‘conservadora’, sempre defendendo a rigidez metodológica e uma forma específica de desenvolver as críticas às fontes (2012, p. 5), é perceptível que essa digitalização da prática histórica (não somente dos arquivos e fontes) é um debate tão extenso quanto difícil. Entretanto, é indispensável fazê-lo enquanto avançamos cada vez mais à virtualização do mundo - por exemplo, no ano que este trabalho é escrito, acontece a tentativa de integralização à sociedade do *Metaverso*, espaço totalmente virtual e digital⁹ onde pessoas, empresas e marcas poderão se encontrar e interagir entre si -, bem como em um momento em que a disputa pela narrativa histórica está diretamente ligada a como utilizamos a internet e a rede mundial de computadores¹⁰.

Pensar a história digitalmente é, fundamentalmente, pensar na inserção dos/as profissionais de história no debate público, nas disputas pelas narrativas e no alcance a novos públicos. Isso, decerto, não é inventada pela digitalização da história (e nem queremos que seja lido dessa forma); De Certeau deixa claro que “da reunião dos documentos à redação do livro, a prática histórica é inteiramente relativa à estrutura da sociedade” (2013, p. 66), ao mesmo tempo que também argumenta sobre a ‘obra de valor’ em história ser “aquela que representa um progresso com relação ao estatuto atual dos “objetos” e dos métodos históricos”, tornando possíveis novas pesquisas (p. 64). Essa inserção do/a historiador/a na estrutura da sociedade, bem como na estrutura acadêmica, é traço consolidado da própria história. Porém, essas mudanças históricas delimitam mudanças metodológicas, portanto, é válido lembrar quais são os espaços ocupados pelos/as historiadores/as tradicionais em comparação com aqueles/as público-digitais, uma vez que espaços diferentes geram demandas diferentes e, assim, novos debates e técnicas¹¹. Cabe estabelecer que, ao que entendemos, as

⁹ Concebido pelo e a partir do antigo *Facebook*, ou seja, essencialmente corporativo, e já parcialmente adotado por empresas (ou suas sedes) brasileiras, como a operadora de telefonia Tim e, recentemente a cervejaria Amstel.

¹⁰ Por exemplo, há empresas privadas e diretamente alinhadas ao revisionismo e negacionismo histórico produzindo conteúdo digital e hipermediático e (re)escrevendo história(s) de forma autônoma, muito alinhadas à extrema direita e/ou ao neoliberalismo. Pensar como combater essas empresas e esses espaços multimilionários de produção de negacionismo é ponto chave ao pensar a HD.

¹¹ Creio que, embora a operação historiográfica, na sua finalidade, tenha uma aproximação clara entre o antes e o agora, esta história dita digital oferece novas demandas e, quando citamos a inserção dos/as seus profissionais no debate público, é inseri-los/as em um novo espaço de atuação. Percebemos esses espaços como diferentes e, por vezes, distantes, entre as diferentes gerações de historiadores/as. Os espaços e as narrativas que historiadores/as “tradicionais” disputam não são os mesmos desta “nova” geração simplesmente porque a atuação daqueles nem sempre se estende ao local do digital, onde estes estão diretamente alocados.

tecnologias digitais e novidades computacionais alinhadas à prática histórica não produzem uma ruptura entre dois momentos da historiografia, mas sim uma diferenciação (talvez uma inclusão ou um acréscimo) na própria teoria e escrita da história. Voltaremos a este debate adiante, mas Serge Noiret o suplementa ao escrever que a história digital “não é feita apenas pela utilização de novas ferramentas digitais que facilitam as velhas práticas. Trata-se também do desenvolvimento de uma relação estreita com as tecnologias suscetíveis em modificar os próprios parâmetros da pesquisa” (2015, p. 33).

A HD está inserida em uma esfera teórica em formação e com debates constantemente sendo atualizados por seus e suas pesquisadores/as. Textos considerados recentes (historicamente, é notável que 10, 20 ou até mesmo 50 anos no mundo nem sempre produzem mudanças sistemáticas observáveis) podem ter demandas datadas - citamos acima a necessidade levantada por Rosenzweig e Cohen (2005) sobre a colaboração entre a história e o XML, um formato superado. Ainda podemos propor esse diálogo entre a historiografia e o XML? Evidentemente que sim, entretanto, substituí-lo pelo formato JSON parece muito mais adequado em 2022.

Nessa lógica, a história digital, bem como a história do tempo presente, costumam esse duplo aspecto, em que investigações sobre ferramentas, inovações tecnológicas e discussões feitas dez anos atrás ou evoluíram, ou foram ultrapassadas, mas, isso não significa que quaisquer trabalhos feitos nesse tempo não sejam cruciais para a consolidação do campo que nós investigamos hoje. Por certo, há uma certa obsolescência em alguns debates gerados pela história digital que é perfeitamente natural e nem um pouco ruim/prejudicial. Um exemplo disso - supracitado - é o debate sobre o uso do XML por historiadores; outros, ainda, comumente datados à primeira década do século XXI, remetem a como a internet e os computadores ainda eram bastante ignorados pelos/as historiadores/as naquele momento. Denis Rolland, em 2004, inicia um texto com a frase “Há uma fonte nova, ainda pouco estudada, em rapidíssima expansão, totalmente inscrita no tempo presente e que deu ensejo a muito poucos exames críticos: a internet” (ROLLAND, 2004, p. 1). Há um certo consenso entre a maioria, se não o todo, dos/as historiadores/as, sobre como o uso da internet não é mais uma dificuldade ou mesmo novidade para os/as profissionais da história. Mas creio que precisamos mesmo enfrentar certas demandas para que possamos substancializar o campo. O custo do avanço teórico é ultrapassarmos essas pendências. Logo, isso não pressupõe a inexistência de conteúdo para ser investigado e estudado em cada uma das bibliografias que contribuíram para a construção da história ou humanidades digitais como são hoje, jamais

reduzindo esses domínios exclusivamente às últimas produções. Por outro lado, essas produções (as atuais) muitas vezes despertam questionamentos e apontamentos sistemáticos (e válidos) sobre o funcionamento da história digital que contribuem para uma estabilização sobre o ela representa(rá), mas que também retornam ao debate acadêmico algumas dessas demandas passadas. Recentemente, em sua tese, Silveira afirma que “apesar de todas as ferramentas computacionais disponíveis, a história digital permanece no horizonte”, tendo as “promessas” e os potenciais da HD sido frustrados pela habituação dos/as historiadores/as com as tecnologias e ferramentas digitais, progressivamente tornando-as objetos que não mais interessam a esta área (2018, p. 37); o autor termina o raciocínio dizendo que o campo “se tornou presa do próprio domínio” (idem). A ideia de a HD ser uma constante “promessa” vem acompanhada de citação para Cameron Blevins, intelectual estadunidense responsável por esse debate, em 2016, que também argumenta que a HD mais contribui para a história pública do que para a “pesquisa que faz argumentos” (utilizo da mesma tradução de SILVEIRA, 2018, p. 36). Evidentemente que não pertence a nós afirmar que Blevins está certo ou errado, porém, cabe a indagação sobre qual é, exatamente, a promessa da história digital? E de onde ela parte, senão de nós mesmos, pesquisadores/as arqueados sobre esse campo tão recente e, reitero, constantemente reavaliado e atualizado?

Sem almejar uma resposta imediata, pensamos que a ênfase é parte processual importante quando definimos quais são as possibilidades que esperamos do campo. Lucchesi (2014, p. 48-49), ao elencar questões a serem discutidas sobre as possibilidades presentes na HD, cita importantes pontos, como as técnicas de *textmining*¹²; iniciativas de trabalhos colaborativos com o público geral; e até mesmo projetos de história digital, como o Identidades do Rio. Cito outros pontos e projetos: a usabilidade das linguagens de programação, como as que citamos neste trabalho (*Python* e *R*, por exemplo), para a produção de materiais historiográficos; o uso de softwares de uso avançado, inteligências artificiais ou *machine learnings*¹³ no processo de historicização da internet e do tempo presente; bem como o excelente *The Trans-Atlantic Slave Trade Database*¹⁴, projeto com um comitê operacional formado quase integralmente de historiadores/as. Penso que a história digital não se propõe a

¹² Traduzido para “mineração textual”, o *textmining* refere-se, basicamente, ao processo de obtenção de dados importantes de um texto. Para os usos do *text mining* nas humanidades digitais, cf. a subseção Humanidades digitais e sociologia computacional. In: Mineração de Texto. Wikipedia. Disponível em: <https://www.wikiwand.com/pt/Minera%C3%A7%C3%A3o_de_texto#/Humanidades_digitais_e_sociologia_computacional>. Acesso em: 20 abr. 2022.

¹³ Dada a complexidade sobre o termo, recomenda-se a leitura sobre *machine learning* disponível em: <<https://tecnoblog.net/responde/machine-learning-ia-o-que-e/>>. Acesso em: 21 mar. 2022.

¹⁴ cf. *Slave Voyages*. Disponível em: <<https://www.slavevoyages.org/voyage/database/>>. Acesso em: 19 mar. 2022.

simplesmente transferir do *kósmos* digital para o acadêmico as novidades e descobertas referentes a ferramentas e tecnologias computacionais. Ao contrário, dedica-se à produção (muito mais do que à reprodução) de produtos e coleções baseados em dados, fontes e técnicas que tão somente a internet e a computação podem comportar. E, como diz Almeida (2011, p. 11), dada a velocidade sob a qual as inovações tecnológicas existem, a palavra de ordem é adaptação.

Ainda assim, compartilho algumas inquietações presentes nos escritos que buscam tentar definir a história digital ou as humanidades digitais, estando eu mesmo me alocando nesta posição de ajudar a lapidar as reflexões através de questionamentos. Nesse sentido, Laitano é sucinto na sua indagação:

Se o digital está em franco processo de sofisticação, e se estamos sempre introduzindo suas novidades em nossas rotinas, como enxergá-lo, no âmbito da história e das ciências humanas, na qualidade de uma reflexão teórico-metodológica? Acredito que, para responder à interrogação, ainda que de forma parcial e provavelmente inconclusiva, é preciso entender, também, as razões pelas quais a chamada “história digital” surge, como uma área do conhecimento, um campo ou um método autônomo, em meio a um outro conceito mais abrangente que, em princípio, já comportaria em si os sentidos que aquele coloca como novidade. Falo das humanidades digitais. (LAITANO, 2021, p. 14)

Falando sobre as humanidades digitais, não se diferencia da história digital que esse conceito também seja um espaço de disputas teóricas sobre o que é e o que faz, embora talvez haja maior consenso nesse sentido. De acordo com Matthew Kirschenbaum (2012, cap. 1), o termo surge em uma espécie de “conversa de boteco”, ou, sob o termo utilizado nas grandes *startups*, em um *brainstorm*, entre John Unsworth, Andrew McNeille e Ray Siemens, sujeitos envolvidos na produção e publicação do livro *Companion to Digital Humanities*, publicado em 2004. Na conversa, Siemens sugere que o livro seja intitulado “Companion to Digitized Humanities” e, para fugir da ênfase na digitalização das humanidades, Unsworth sugere o que, de fato, se tornou o título do livro. Unsworth tinha, portanto, a vontade de demonstrar a ideia de uma área, as humanidades, que fosse digital, não digitalizada. Sob outra ótica, antes de ser *digital humanities*, chamavam esse inconcluso e amplo campo de *humanities computing*, associando-o exclusivamente à junção humanistas + computador. Quase vinte anos depois, ainda há muito o que debater, descobrir e reavaliar sobre, mas, de acordo com a Wikipédia (em inglês, pois, curiosamente, não há um verbete em português para o termo),

Humanidades digitais incorporam tanto materiais digitalizados quanto nascidos digitais e combina as metodologias de disciplinas humanísticas tradicionais (como retórica, história, filosofia, linguística, literatura, arte, arqueologia, música e estudos culturais) e ciências sociais, com ferramentas da computação (como hipertexto,

hipermídia, visualização de dados, recuperação de informação, mineração de dados, estatísticas, mineração de texto, mapeamento digital) e publicação digital.¹⁵ (WIKIPÉDIA, tradução nossa)

E talvez esse seja um lacônico resumo sobre o que abrange as humanidades digitais. Para além disso, entretanto, precisamos revisitar as humanidades digitais como uma área em franca expansão que, talvez, não precise ser uma área; pode ser uma disciplina, um espaço, uma forma de intervenção da/na academia ou uma comunidade. Daniel Alves (2017) diz que

assumir as Humanidades Digitais simplesmente como uma comunidade, sem preconceitos ou sentimentos de inferioridade, em vez de perseguir o sonho da institucionalização, pode ser um caminho para consolidar os contributos positivos que têm resultado da incorporação e da reflexão sobre o digital nas Humanidades e na sociedade em geral. (ALVES, 2017, documento *online*)

A ideia das humanidades digitais enquanto comunidade, pelo menos para uma breve introdução sobre o que significa esse conceito emergente, é agradável quando assumindo a característica básica dela: a intercomunicação entre diversas áreas do saber - acadêmicas ou não -, desde história até as ciências computacionais (falaremos melhor sobre a multidisciplinaridade no próximo capítulo). Dessa forma, me parece justo ocasionalmente citar as humanidades digitais como um espaço comunitário, de partilha e divisão de tarefas para a produção de um projeto maior. É um local de trocas. Alves continua, dizendo que o “recurso ao conceito de «comunidade» funcionou e funciona ainda como uma forma de dar coerência ou consistência, de criar uma identidade entre investigadores que vêm de áreas disciplinares muito díspares ou de disciplinas que, aparentemente, pouco comunicam entre si” (2017, documento *online*). Essa identidade é muito importante, uma vez que essa diferença disciplinar pode ser motivo de um natural processo de distanciamento, afinal, o que as ciências da informação precisam da historiografia e vice-versa? Transformar essa interação em uma comunidade é a maneira mais inteligente de fazer a manutenção dessas redes de relacionamento que nos são tão caras. Em outro texto, Stephen Robertson, que inicia escrevendo que as humanidades digitais são como uma “grande barraca” (*big tent*, na expressão original), na sequência argumenta que

seríamos melhor servidos se reimaginássemos as humanidades digitais não como uma grande barraca, mas como uma casa com muitos cômodos, espaços diferentes para disciplinas que são pontos de entrada e corredores para espaços centrais onde

¹⁵ Originalmente “Digital humanities incorporates both digitized (remediated) and born-digital materials and combines the methodologies from traditional humanities disciplines (such as rhetoric, history, philosophy, linguistics, literature, art, archaeology, music, and cultural studies) and social sciences,[6] with tools provided by computing (such as hypertext, hypermedia, data visualisation, information retrieval, data mining, statistics, text mining, digital mapping), and digital publishing”

aqueles de disciplinas diferentes, trabalhando com ferramentas e mídias específicas, podem se reunir. (ROBERTSON, 2016, cap. 25, tradução nossa)¹⁶

Talvez Robertson tenha captado a essência mais simples das humanidades digitais de forma a incorporar essa ótima analogia no seu texto e fazer o leitor realmente imaginar um grande salão comunal onde todas as ‘disciplinas se reúnem’ e discorrem sobre a utilização comunitária de suas próprias ferramentas de trabalho. As palavras de ordem, dessa vez, são compartilhamento e rede de relacionamento (ou, *networking*). As humanidades digitais são uma sociedade do conhecimento e da pedagogia em consonância com a publicização da história em um espaço colaborativo entre nós e eles - nós enquanto os/as profissionais de história **ou** os acadêmicos em geral e, eles, enquanto os/as profissionais de outras áreas do conhecimento **ou** o grande público, por vezes não-especialista. Além de comunidade, Laitano (2021, p. 21) também encara as humanidades digitais como um movimento ou uma causa à qual pesquisadores/as podem aderir, a exemplo de um movimento social que decidimos defender. Já Damian e Almeida falam que ela "surge como um campo interdisciplinar disposto a dar guarida às reflexões e às práticas suscitadas pelas mudanças decorrentes da introdução das tecnologias digitais no universo da cultura" (2015, p. 8).

Por outro lado, percebo as humanidades digitais como a representação de oportunidades. São as oportunidades de produzir um projeto que, por vezes, seria enclausurado dentro do próprio departamento de história das universidades; ou de dividir a nossa práxis científica com outrem e apreender a deles; ou então, a oportunidade de perceber o grande público como agente central na produção que almejamos. São trocas de interesses que possibilitam a realização de iniciativas que outrora jamais seriam plausíveis pensarmos e que, na era digital, são muito valiosas. É importante, para tanto, que esse espaço seja constantemente revitalizado e que tenhamos diversos campos da historiografia inseridos aqui; que sejam as humanidades digitais uma grande casa que acomoda não apenas a história digital, a história pública e a história oral, tampouco somente o tempo presente.

Entendendo essas percepções, percebe-se o quanto os historiadores e os humanistas digitais são semelhantes, principalmente ao perceber como uma grande parte desses projetos dos humanistas digitais são, para fins de comparação, projetos históricos. Não só isso, mas os

¹⁶ Originalmente, “we would be better served by reimagining digital humanities not as single all-encompassing tent but as a house with many rooms, different spaces for disciplines that are not silos but entry points and conduits to central spaces where those from different disciplines working with particular tools and media can gather”

desafios são compartilhados também, visto que ambos precisam enfrentar os ônus que a internet traz à pesquisa científica de dados: a velocidade com que a geração de dados ocorre devido ao *Big Data* (termo introduzido na introdução deste trabalho e que será revisitado nos próximos capítulos); a alta demanda gerada pelas redes sociais computacionais (interações imediatas via comentários, *hashtags* ou assuntos em alta no *Twitter*, *Instagram*, *Youtube*, etc.); a dificuldade em usar e ampliar o uso de *softwares* de uso avançado ou linguagens de programação. São dificuldades compartilhadas que fazem o crescimento e aderência desses e a esses campos serem mútuos. Mas, como podemos pensar a escrita da história neste íterim? Podemos pensar em uma ruptura da metodologia de antes com agora? A *obra de valor* descrita por De Certeau mantém-se na era digital? Diante das reflexões colecionadas nos últimos parágrafos deste trabalho, pensamos que parte desses debates podem ser identificados.

A escrita da história é, para todos os fins, feita sob certo método científico de bastante rigor, mas **como** ela é feita difere-se de outro momento. Reiteramos não ser tão somente a mera atividade de escrever que mudou (do analógico para o digital), mas a forma como apreendemos as fontes, fazemos nossas críticas, produzimos nossos dados e acrescentamos o nosso trabalho com outras mídias que, sim, mudou muito. Poder investigar nos arquivos digitalizados, nos repositórios disponíveis *online*, nos PDFs, é um passo muito importante otimizado pela rede mundial de computadores, mas, investigar nos arquivos nascidos digitais, nos vídeos (exclusivamente) disponíveis no *YouTube* (como Laitano fez), nas iniciativas criadas, consolidadas e disponíveis somente na *internet* - todas possíveis fontes primárias para nossas pesquisas -, é, talvez, o mais importante espaço que a história digital construiu. Existe, inclusive, uma certa presença democrática a ser discutida na natureza dessas coleções nascidas-digitais, mesmo sob a perceptível desigualdade, onde muitas pessoas não têm acesso a um computador ou à internet. Bolick, em 2006, fala que

A criação de arquivos digitais alterou a dinâmica de fazer pesquisa histórica ao mudar quem é capaz de conduzir a pesquisa e como a pesquisa histórica é feita. **Arquivos digitais são coleções de dados numéricos, textos, imagens, mapas, vídeos e arquivos de áudio** que estão disponíveis na Internet (BOLICK, 2006, p. 122, tradução e grifo nossos)¹⁷

¹⁷ Originalmente “The creation of digital archives has shifted the dynamics of doing historical research by changing who is able to conduct the research and how historical research is done. Digital archives are collections of numerical data, texts, images, maps, videos, and audio files that are available through the Internet.”

Com o avanço tecnológico, arquivos digitais podem, também, conter bancos de dados virtuais (que podem ser gerados de forma automatizada por *bots*¹⁸, como o Buppa¹⁹, ou “manualmente”, através da programação), projetos de autoridade compartilhada (*crowdsourcing*) - projetos que podem, inclusive, **criar repositórios** próprios -, e, talvez um dos quesitos mais interessantes a serem tocados atualmente, **metadados**, os “dados sobre outros dados”, gerados por inteligências artificiais, *bots*, códigos de programação criados para minerar esses dados ou algoritmos. Esses metadados são capazes de gerar infinitas interpretações e visualizações dos próprios dados, de comunidades inteiras, termos, usuários, assuntos, interações. Esses dados se tornam praticamente universais uma vez que atingem a *web* e, com as ferramentas certas, são passíveis de serem coletados, ao passo que, com o método, a crítica e o fazer científico certos, são naturalmente *historicizáveis*. Conseqüentemente, a infinitude de janelas interpretativas criada por esses arquivos digitais altera a forma como podemos enxergar, entender e produzir teoria da história, uma vez que a própria dinâmica das sociedades de conhecimento (*scholarships*) da historiografia é afetada. Para tanto, os questionamentos que ficam são: de qual(is) forma(s) podemos pensar a escrita da história hoje? Se não exclusivamente textualmente ou através de gráficos formulados no *Microsoft Office*, podem nossos trabalhos serem apresentados com outras características? E quais usos podemos fazer das ferramentas computacionais para nos auxiliar - i.e., podemos, por exemplo, agregar nossas pesquisas com o uso de grafos produzidos com dados da internet em *softwares* de uso avançado? Procurarei trilhar alguns caminhos para responder essa questão no capítulo três, explorando essas ferramentas e esses formatos de dados.

Sem dúvida, muito disso passa pelo que já foi escrito aqui, logo, por essas mudanças observadas nos últimos anos, e não só nos produtos resultados pela utilização das ferramentas computacionais, mas com as aproximações que podemos fazer através das humanidades digitais - leia-se, da multidisciplinaridade presente nesse contexto.

3. “MULTIDISCIPLINARIDADE DIGITAL”: as linguagens de programação e a ciência computacional agregadas às humanidades digitais.

As Humanidades Digitais também possibilitam uma visão crítica do uso e apropriação tecnológica, a partir da qual é importante exigir e observar como as tecnologias digitais devem ser cada vez mais adaptativas e adequadas a cada contexto específico onde são aplicadas e buscam alternativas para gerar autonomia e

¹⁸ *Bot* é a denominação para os softwares programados para fazerem aplicações padrozinadas, como um robô faria. Dessa forma, um *bot*, na internet, reproduz continuamente uma tarefa programada.

¹⁹ Exemplo de *bot* do *Twitter* criado pelo Laboratório de Estudos sobre os Usos Políticos do Passado (LUPPA). Sua atuação pode ser conferida pelo seu perfil na rede social. Disponível em: <<https://twitter.com/buppabot>>

resistência às hegemonias que promovem a dependência comercial e promover a cultura livre para incentivo e o letramento digital que viabilizam o bem comum. (LONDOÑO, 2020²⁰)

Sob a alcunha multidisciplinar, as humanidades digitais perpassam por esses diferentes saberes, aqueles cômodos de uma casa que acomoda tantas disciplinas. Pensando nessa metáfora, talvez o que está sob as cabeças de todas as disciplinas, o telhado de toda a casa, é o que nós chamamos de *big data*. Esses *grandes dados* que estão indefinidamente distribuídos por todo o universo digital, principalmente a partir da década de 2010, hoje são definidores do comportamento da internet mundial, desde o usuário mais singelo até a corporação mais colossal. Todos nós, todos eles, somos controlados e controlamos uma série de algoritmos que consolidam essa cultura digital de hoje. BOYD e CRAWFORD (2012, p. 663 apud SANTAELLA, 2021, p. 7), definem três pontos-chave para o entendimento do *big data*:

1. Tecnologia: maximização do poder computacional e da precisão algorítmica para juntar, analisar, combinar e comparar grandes conjuntos de dados;
2. Análise: desenho de grandes conjuntos de dados para identificar padrões a fim de responder a demandas econômicas, sociais, técnicas e jurídicas;
3. Mitologia: a crença difundida de que grandes conjuntos de dados oferecem uma forma mais elevada de inteligência e conhecimento que gera insights previamente impossíveis, envolvidos na aura da verdade, objetividade e precisão.

Embora não seja recente e, certamente, hoje possa ser reestruturada, essa definição é suficiente para um entendimento geral sobre o espaço ocupado pelo *big data*. Talvez para este trabalho os itens 1 e 2 sejam os mais importantes, uma vez que é justamente isto que pretendemos demonstrar aqui, em conjunto com a própria prática histórica. As humanidades digitais fazem, portanto, essa intermediação para que as ciências da informação possam auxiliar a história em captar, interpretar e visualizar esses *grandes dados*, ao passo que nós, as ciências humanas, os auxiliamos com perspectivas humanísticas das suas pautas e com a rigidez e especificidade científica com que tratamos as fontes.

A utilização do *big data* se tornou anexa ao campo das humanidades digitais e, seu entendimento, impreterível para os/as pesquisadores/as do campo, fazendo com que, com os algoritmos e o *big data*, “as humanidades abriam-se para a observação de problemáticas complexas, ligadas a questões que afetam o bem estar social.” (SANTAELLA, 2021, p. 10).

Hoje, portanto, as possibilidades que a história tem graças às humanidades digitais e essa multidisciplinaridade cujos caminhos construídos nos levaram até o encontro com as ciências computacionais e informacionais, bem como com a programação, são muito mais amplas. O uso de *softwares* de código aberto (cujo uso não está ligado a uma empresa ou

²⁰ In: LONDOÑO, Felipe. Entrevista concedida a Isabel Jungk. **TECCOGS – Revista Digital de Tecnologias Cognitivas**, n. 21, jan./jun. 2020, p. 12-23

pessoa e, portanto, não pode ser descontinuado sem a nossa aprovação), a produção de jogos eletrônicos enquanto ferramentas educacionais, mapas de georreferenciamento, cenários de realidade aumentada (como a produção de uma museologia transmídia), os *bots* supracitados neste trabalho, a mineração e interpretação de dados que manualmente seria impossível fazer e a produção do conhecimento na nuvem são facilidades trazidas pelo entrelace dessas áreas. No Brasil, temos projetos incríveis que estão em constante uso desses produtos, como o *Caminhos da Ditadura em Porto Alegre*, projeto produzido pela pesquisadora Anita Natividade Carneiro, em colaboração com o LUPPA/UFRGS²¹; ou então o Laboratório de Estudos sobre Imagem e Cibercultura (LABIC), que recorrentemente produz grafos visuais (sobre os quais falaremos melhor no próximo capítulo deste trabalho) sobre temas “em alta” na redes sociais²²; ou o *tour virtual* que ocorre no site do Museu do Amanhã, onde pode-se fazer um passeio totalmente imersivo e interativo pelo Museu diretamente de um aparelho eletrônico²³.

Todos esses projetos e muitos outros que existem no mundo são imersões tecnológicas que existem pela transdisciplinaridade entre diferentes áreas do saber, inclusive a história, a museologia, a arquivologia, entre outros campos humanistas. A atuação conjunta dos conhecimentos adquiridos na historiografia e na computação permite que esses projetos existam e que, portanto, sejam disponibilizados aos públicos. As humanidades digitais fazem, constantemente, produtos de história pública também.

Existe, nesse sentido, uma intercambialidade necessária ao avanço que temos tido nos últimos anos, tanto na nossa área quanto no mundo tecnológico como um todo. É bidirecional (ou multidirecional) essa troca de conhecimentos entre os campos, mas, mais do que isso, é multidisciplinar o suficiente para que algumas concretudes históricas sejam revisitadas e repensadas. Talvez, nesse sentido, o conceito de universidade precise de revisão urgente, uma vez que sabemos que a multidisciplinaridade não ocorre mais pontualmente, em alguns casos, mas praticamente em todo o cenário acadêmico-escolar. Desde o uso de aparelhos eletrônicos em sala de aula até a possibilidade de aprendermos a utilizar técnicas que não são “nossas” para atividades acadêmicas, se tornaram necessidades básicas à formação do/a cientista

²¹ cf. Mapa localiza espaços de violações dos direitos humanos durante a ditadura militar em Porto Alegre. Disponível em: <<https://gauchazh.clicrbs.com.br/porto-alegre/noticia/2022/02/mapa-localiza-espacos-de-violacoes-dos-direitos-humanos-durante-a-ditadura-militar-em-porto-alegre-ckz7409f1003z015pfsqa59x.html>>

²² cf. LABIC. Disponível em: <<https://www.labic.net/>>

²³ cf. Museu do Amanhã - tour virtual 360°. Disponível em: <<https://museudoamanha.org.br/tourvirtualpratodomundo/>>

humanista²⁴. A existência de laboratórios e centros de estudos que se pautam sob essa multidisciplinaridade sistemática é importante, mas ainda não exclui a necessidade de repensar e reestruturar a universidade como um todo para que se torne um local com uma infraestrutura e suporte suficientes para a existência de disciplinas, pós-graduações e especializações que levem em consideração esse intercâmbio. É pensar não somente na existência de profissionais aptos ao mercado tradicional de suas áreas, mas às novas exigências trazidas pelas próprias mudanças governamentais sobre o mercado de trabalho.

Embora também haja certos debates sobre o impacto dessa hipermedialidade circunscrita à multidisciplinaridade em valores que não pertencem a este trabalho (os impactos ambientais, sociais, políticos e econômicos), nos ocorre citá-los, a fim de esclarecer que não nos interessa esgotar quaisquer discussões sobre o campo, tampouco reduzi-las ao que pretendemos aqui trazer.

Sobre essa integralização da história, das humanidades e das tecnologias da informação, é importante ressaltar que os ganhos que os profissionais das ciências humanas têm são engrandecedores. Seus e suas pesquisadores/as podem, cada vez com maior facilidade, adaptarem seus trabalhos ou fabricarem novas pesquisas com ferramentas que facilitam a própria conciliação entre os fazeres acadêmicos e a publicização da história, ou mesmo a simplificação de problemas, dados e interpretações que, por vezes, são complexas. Não é nossa intenção romantizar o uso das ferramentas digitais para as ciências humanas, uma vez que, sim, são produtos laboriosos e de difícil introdução no campo teórico das nossas pesquisas. Mas é inerente a este trabalho desmistificar certas idealizações que podem surgir em torno da ideia de utilizarmos, por exemplo, linguagens de programação, códigos, etc.

Penso em fazê-lo partindo da própria práxis histórica, afinal, *o que faz o/a historiador/a?* Sem a intenção de aprofundar o debate ou de remeter às discussões já realizadas por Koselleck, De Certeau e Hartog, partimos do básico: produz, escreve, cria e interpreta história e narrativas, passados e presentes. Tudo isso a partir de uma rigidez científica e metodológica específica, como já citado neste trabalho. Pois bem, *o que faz o/a programador/a?* Longe de querer esgotar essa dúvida com uma única resposta, me aventuro em utilizar as palavras de Leandro Pinho Monteiro, que escreve que “Programar, na prática, é

²⁴ A abordagem que faço pode passar pela noção de *(in)disciplina da história*, corrente trabalhada por historiadores brasileiros nos últimos anos, sob a ideia de repensar os fundamentos epistemológicos da historiografia e reorganizar a parte curricular da história. Para uma leitura mais aprofundada da temática, cf. ÁVILA, A. L. de. O fim da história e o fardo da temporalidade. Tempo e Argumento, Florianópolis, v. 10, n. 25, p. 243 - 266, 2018. Acesso em: 22 abr. 2022.

escrever um texto que será transformado em um software”²⁵ - com uma cientificidade própria. Enquanto historiadores-programadores/as, escrevemos, portanto. Escrevemos códigos textuais que serão validados e utilizados para a escrita dos nossos artigos, projetos e trabalhos acadêmicos-escolares. Escrevemos, de certa forma, sob novos “idiomas” (mesmo porque universalmente os códigos, além de serem escritos em linguagens que não-faladas, também são escritos em inglês), o que nos leva à necessidade de apropriação deles.

Simplificadamente, as linguagens de programação são conjuntos de regras e métodos que criam códigos-fonte que, quando ‘lidos’ por um dispositivo, possuem instruções específicas para ações que este dispositivo executará. O código-fonte é uma forma lógica de leitura que pode ser transformado em um *software* ou servir como base de interpretação do dispositivo. Dessa forma, essas linguagens, quando escritas seguindo as regras e métodos dispostos a elas, criam um documento que será executado e retornará um resultado/produto. Há alguns anos, no auge dos *blogs*, muitas pessoas que utilizavam plataformas como o *Blogspot* ou *Wordpress* tinham contato com uma linguagem básica chamada HTML (*HyperText Markup Language*), que compunha o código responsável pelo visual gráfico do *blog*, fazendo com que o/a usuário/a, por vezes, se aventurasse a tentar entender e alterar algumas linhas deste código-fonte do *blog* a fim de modificar a estrutura gráfica do site (como botões, banners, cores, etc.). Esse contato era feito inclusive por usuários/as comuns, não-programadores/as, dispostos/as a minimamente entender a lógica por trás do código.

Pois, no caso desta pesquisa, trabalhamos com uma linguagem de programação chamada *Python*²⁶, cujos códigos escritos nela são chamados de *scripts* e são executados na própria máquina, contanto que esta disponha dos pré-requisitos para a execução (a depender do *script* e da finalidade que pretendemos com o nosso código). Uma das suas principais características (que facilita seu uso por pessoas não especializadas) é a possibilidade de um código com poucas linhas já ser suficiente para executar um processo específico. Para tê-lo e utilizá-lo, basta ao usuário fazer o *download* da última versão do *Python* pelo próprio site e instalá-lo em um computador. A partir daí, existem “raízes” conhecidas como as bibliotecas padrão do *Python*, responsáveis por expandir a capacidade de atuação e produzir uma linguagem com mais funcionalidades e possibilidades. Falando sobre essas bibliotecas, Arnold e Tilton argumentam que

²⁵ cf. MONTEIRO, L. P. **O que é linguagem de programação?** Universidade da Tecnologia. s. d. Disponível em: <<https://universidadedatecnologia.com.br/o-que-e-linguagem-de-programacao/>>. Acesso em: 08 de abril 2022.

²⁶ cf. Python. Disponível em: <<https://www.python.org/>>. Acesso em: 08 de abril 2022

Todas essas bibliotecas operam nas mesmas estruturas básicas de dados internos, permitindo uma fácil interoperabilidade entre as bibliotecas. Estatísticos podem manter flexibilidade enquanto conduzem a exploração de dados como um resultado direto dessas bibliotecas, que operam com uma linguagem de programação única, em vez de ficarem utilizando ferramentas desconexas (uma das outras). (ARNOLD & TILTON, 2019, Cap. 24, tradução nossa²⁷)

Para as humanidades digitais, há também uma outra linguagem bastante conhecida²⁸, principalmente por aqueles/as que trabalham com a gestão de ferramentas voltadas às estatísticas e/ou ciência de dados, conhecida como *R*²⁹.

Para este trabalho, nosso destaque ficará sobre o *Python*, em específico sobre uma biblioteca específica conhecida como *Twint (Twitter Intelligence Tool)*³⁰. A partir do momento que o/a pesquisador/a torna-se minimamente familiar com a linguagem, a instalação da biblioteca se torna fácil, principalmente visto que há um tutorial disponível na sua página do *GitHub*. O *Twint* serve como uma ferramenta, escrita e executada em *Python*, que faz uma raspagem e coleta de dados do *Twitter*, a rede social virtual em forma de microblog muito utilizada e famosa até hoje (e talvez uma das mais longevas da história). A partir de um código básico, a ferramenta rapidamente faz uma coleta de postagens que contenham uma palavra-chave que o usuário escolher, na data que escolher. Pode ser um termo, um usuário/perfil, uma *#hashtag*, etc. Essa raspagem vira um arquivo que pode ser visualizado enquanto tabela (aberto via *Excel* ou *Google Sheets*, por exemplo), contendo diversas informações, como data e hora, usuário que publicou, mensagem publicada, possíveis mídias contidas na mensagem, número de curtidas e compartilhamentos, etc.

Esses dados, na sua forma bruta, já podem ser utilizados por um/a intelectual disposto/a a manualmente (i.e., um a um) interpretá-los, mas também podem ser tratados e interpretados de outras formas, através de outros dispositivos (automatizadamente, diga-se de passagem). Uma delas, a que constrói a idealização deste projeto, é através de grafos³¹. Há

²⁷ Originalmente, “*These libraries all operate on the same basic internal data structures, allowing for easy interoperability between libraries. Statisticians are able to maintain flexibility while conducting data exploration as a direct result of these libraries that operate with a single programming language, rather than using disconnected one-off tools.*”

²⁸ Para um entendimento do uso prático do *R*, sugiro a leitura de duas lições postadas no site *Programming Historian*, uma já traduzida para o português e outra em inglês. Cf. ARNOLD, T. & TILTON, L. **Processamento básico de Texto em R**. *Programming Historian*. Disponível em: <<https://programminghistorian.org/pt/licoes/processamento-basico-texto-r>> e BREY, A. **Temporal Network Analysis with R**. *Programming Historian*. Disponível em: <<https://programminghistorian.org/en/lessons/temporal-network-analysis-with-r>>. Acesso em 8 abr. 2022.

²⁹ Para uma concepção mais ampla sobre o *Python* e *R*, cf. ROCHA., R. **R vs Python – Uma Análise Desapaixonada**. FLAI. Disponível em: <<https://www.flai.com.br/ricardo/r-vs-python-uma-analise-desapaixonada/>>. Acesso em: 08 de abril 2022.

³⁰ cf. *twintproject/twint*. *GitHub*. Disponível em: <<https://github.com/twintproject/twint>>. Acesso em: 08 de abril 2022.

³¹ Voltarei ao conceito e explicação dos grafos, bem como trabalharei melhor com eles no [próximo capítulo](#) deste trabalho.

outras, como tabelas com pontos específicos dessa coleção, gráficos (que não devem ser confundidos com grafos), análise de discurso, mapas conceituais, entre outros. E a própria linguagem *Python* também fornece outras bibliotecas que podem conjuntamente ao *Twint* produzir especificidades interessantes, como a análise de sentimentos³².

Sem dúvida, é um universo tão extenso quanto pode-se pensá-lo, ao passo que sua utilidade é, para as ciências humanas, enriquecedora. Enquanto esta pesquisa pensa em um caminho específico a ser seguido, outras formas de utilização da programação, de programas de código aberto e de modelos de visualização e interpretação de dados podem surgir futuramente. Tem espaço, campo de atuação e modelos de negócio para todos/as que estejam propensos/as a produzir história com a internet. Esta é, para todos os fins, praticamente infinita, existindo diversos conglomerados virtuais que são potenciais espaços de análise do/a humanista e historiador/a e, inclusive, potenciais fontes (nascidas digitais, principalmente). Portanto, para além do que esta pesquisa oferece e pretende elucidar sobre a operação histórica no universo hipermediático e digital da internet, há outros locais inexplorados que, dentro do que entendemos por história digital, são produtos inexoráveis que servem à historiografia como auxiliares diretos do e no nosso trabalho.

Para além disso, com a ciência da computação em consonância ao nosso trabalho, há a possibilidade de nos inserirmos em novas narrativas e espaços culturais e historicizarmos problemas sociopolíticos muitas vezes inacessíveis a nós, como expressa Felipe Londoño:

As redes, utilizadas de maneira inteligente, dão voz a muitas comunidades que tradicionalmente não têm conseguido expressar suas opiniões sobre temas de interesse comum e encontrar canais de comunicação alternativos de visibilidade e busca de soluções aos problemas e conflitos dos grupos humanos. [...] por meio dos fluxos de informação que transitam em suas redes, é possível detectar movimentos sociais e manifestações políticas, bem como comunidades virtuais e práticas que apresentam outras formas de compreender os mundos que habitamos. (LONDOÑO, 2020, p. 20)

Este alcance a realidades de difícil acesso não é novidade exclusiva da digitalização das humanidades, mas trata-se de uma facilitação que pode-se obter a partir da intercomunicação entre a história, a internet e a ciência de dados, por exemplo.

O escopo que pretendemos abrir com estas colocações colabora com a ideia apresentada no capítulo anterior de quais são os almejos da história digital e por quê ela não é meramente a agente de digitalização da academia tradicional. Essas incorporações reservadas à união do nosso cômodo com outros cômodos da grande barraca das humanidades digitais

³² Para um maior entendimento sobre o que se trata, cf. Análise de Sentimentos em Python. Disponível em: <<https://medium.com/@alegeorgelustosa/an%C3%A1lise-de-sentimentos-em-python-2a7d04a836e0>>. Acesso em: 08 de abril de 2022.

promovem a capacitação de pesquisas sob novas perspectivas teórico-metodológicas, sem separá-las do mérito histórico-acadêmico que é o método científico.

No que concerne à *história vista de baixo*³³, Thompson, Christopher Hill e Natalie Davis provavelmente ficariam contentes com os caminhos que estão sendo tomados com e a partir das humanidades digitais. Com a série de publicações intituladas *Debates in the Digital Humanities*, desde o início da década de 2010 é possível acompanhar artigos sobre diversos temas emergentes dentro das humanidades digitais. Entre estes artigos, há publicações sobre o uso das humanidades digitais no estudo da cultura negra, do feminismo, da história da escravidão e, inclusive, tendo organizações textuais específicas para essas temáticas sociais. Em 2019, fora lançado o livro digital *Bodies of Information*, sob organização de Jacqueline Wernimont e Elizabeth Losh³⁴, carregando artigos que relacionam as questões de gênero, feminismos interseccionais e humanidades digitais e, em 2021, o *The Digital Black Atlantic*, organizado por Roopika Risam e Kelly Baker Josephs³⁵, coleciona escritos sobre as humanidades digitais em comunicação com os estudos sobre o *black atlantic* (termo cunhado por Paul Gilroy para falar sobre a fusão da cultura negra com outras culturas ao redor do atlântico).

Para além de produções acadêmicas que respondem constantemente (i.e., bimensalmente, semestralmente, quinzenalmente, a depender da organização de cada revista, editora, etc.) a pergunta “o quê está sendo feito nas humanidades digitais no mundo?”, esses livros e editoriais também carregam a qualidade de estarem produzindo as fontes nascidas digitais e, mais ainda, de nos mostrar *outras formas* de produzi-las. Digo isso porque, enquanto escrevo este trabalho e demonstro como o *Twitter* e os grafos podem ser tratados como fontes nascidas digitais, me encontro com maneiras em que outras ferramentas auxiliam nesse aspecto. Fosse tão somente as redes sociais virtuais e a manipulação de dados que produzissem fontes nascidas digitais não-textuais, seria mais simples. Entretanto, as estatísticas, as artes, projetos envolvendo ideias antes inimagináveis, são espaços performáticos para fazermos uma reavaliação da nossa própria posição enquanto pesquisadores/as humanistas e não esquecermos jamais que existem fronteiras invisíveis que

³³ A história vista de baixo ou a história popular é uma corrente historiográfica datada do século XX, na Inglaterra. E. P. Thompson, Natalie Zemon Davis e Christopher Hill são expoentes importantes da corrente. A partir dela, os/as historiadores/as pretendem a historicização de grupos invisibilizados pela historiografia tradicional, como as mulheres, os/as escravizados/as, os/as negros/as, entre outros.

³⁴ Para maiores e mais profundas informações, cf. LOSH, E. & WERNIMONT, J. (eds.) **Bodies of Information - Intersectional Feminism and Digital Humanities**. Minneapolis, MN: University of Minnesota Press, 2019. E-book. Disponível em: <<https://dhdebates.gc.cuny.edu/projects/bodies-of-information>>

³⁵ Para maiores e mais profundas informações, cf. RISAM, R. & JOSEPHS, K. B. (eds.) **The Digital Black Atlantic**. Minneapolis, MN: University of Minnesota Press, 2021. E-book. Disponível em: <<https://dhdebates.gc.cuny.edu/projects/the-digital-black-atlantic>>

precisamos dismantelar no intuito de ampliar o nosso horizonte de expectativas³⁶ sobre a atuação das humanidades digitais.

Certamente muito disso está relacionado ao quão recente é a produção historiográfica na e através da internet - nossa percepção sobre seus usos e suas possibilidades ainda há de ser expandida. A *web semântica* por si só - e em detrimento da *web 2.0* - é recente, datada de 2001. Esta cooperatividade entre humanos e computadores é moderna, logo, nossa caixa de percepção ainda é apertada. De qualquer modo, fará parte do lugar de intelecto de todo/a aquele/a que se dispôr a estudar essa temática constantemente refletir sobre as limitações que nos são impostas pelo desconhecimento de outras técnicas, ferramentas e, principalmente, a desconexão entre as ciências humanas e tantas outras áreas. Localizar os usos da história em conciliação com outros campos é, portanto, uma inquietude importante no fazer histórico deste/a pesquisador/a.

Essa atualização do/a próprio/a pesquisador/a serve à história não apenas pela prerrogativa de estar atualizado/a (ou de “descobrir uma novidade”), mas como forma de se preparar tecnicamente para localizar a própria pesquisa na estrutura social corrente. De Certeau, na sua metáfora, diz que “Como veículo saído de uma fábrica, o estudo histórico está muito mais ligado ao complexo de uma fabricação específica e coletiva do que ao estatuto de efeito de uma filosofia pessoal ou à ressurgência de uma "realidade" passada. É o produto de um lugar.” (DE CERTEAU, 2013, p. 64) ao passo que, adiante, acrescenta que “Se é verdade que a organização da história é relativa a um lugar e a um tempo, isto ocorre, inicialmente, por causa de suas técnicas de produção” (idem, p. 70). A partir desta argumentação, penso que localizar a pesquisa nos diferentes cenários relativos ao universo hipermidiático da internet também é uma forma de localizá-la nesta teia social (neste caso, social virtual) e no próprio tempo (afinal, como fora supracitado neste trabalho, as próprias tecnologias são suplantadas por suas versões atualizadas, trazendo uma certa efemeridade ao tempo histórico na internet).

³⁶ Embora a forma com que eu me aproprio do termo de Reinhart Koselleck não tenha o imperativo de reproduzir fielmente a ideia do autor (principalmente por este não tê-lo feito pensando no ambiente digital como objeto de análise), ainda assim, o horizonte de expectativa que reproduzo perpassa pela meta-história e pela estrutura do conhecimento tanto quanto originalmente. A experiência que paulatinamente surge no horizonte é tanto um espaço do analógico quanto do digital.

4. DADOS DA WEB SEMÂNTICA, SUA VISUALIZAÇÃO ATRAVÉS DE GRAFOS E AS FONTES NASCIDAS DIGITAIS.

O mundo multiforme do acesso livre ao conhecimento por meios digitais (open access), apoiado nas mídias sociais e nas aplicações para celulares, permitiu compartilhar globalmente – e reviver no presente – a história em público. Alcançar universalmente diversos indivíduos e grupos, e compartilhar as experiências históricas do passado, nunca foi tão fácil e à disposição de quem quer que seja. (NOIRET, 2015, p. 45)

O termo *semântico* significa a relação entre conjuntos linguísticos. A *web semântica*, também chamada pela pesquisadora Anita Lucchesi (2014, p. 47) de uma terceira geração da *web* (isso é, posterior à *web 2.0*), surgiu em 2001 e representa a relação entre o homem e o computador; a atuação humana e aquela produzida pela máquina. A priori, essa relação é conectada pelos dados originados pelo usuário em consonância com a máquina e sua finalidade é transformar a rede mundial de computadores em um grande algoritmo capaz de ser lido por uma máquina.

Segundo a *World Wide Web Consortium (W3C)*,

O termo “Web Semântica” refere-se à visão do W3C da Web dos Dados Conectados. A Web Semântica dá às pessoas a capacidade de criarem repositórios de dados na Web, construir vocabulários e escreverem regras para interoperarem com esses dados. A linkagem de dados é possível com tecnologias como RDF, SPARQL, OWL, SKOS. (W3C, documento online)³⁷

Lucchesi (2014, p. 51) observa que essa escrita semântica da história pelos hipertextos, feita no ciberespaço, modifica os próprios elementos básicos de escrita da história. Para mim, essa *semanticidade* da *web* também carrega, para os fins deste trabalho e da operação historiográfica, novas técnicas de produção. A operação historiográfica de De Certeau, portanto, sofre certa modificação pelas próprias técnicas nascidas e regidas pela *web* (e principalmente pela sua versão semântica). A existência do trabalho hipertextual e da história digital carrega a técnica que não mais se estende tão somente ao lugar social, às práticas científicas e à escrita, passando agora por uma hipermedialização multidisciplinar, que terá que ser composta por elementos externos aos da prática histórica *per se*.

A “nova versão da *web*” carrega, portanto, uma certa responsabilidade na própria existência das humanidades digitais (pode ser coincidência, mas ambos os conceitos, *semantic web* e *digital humanities*, surgem em períodos temporais muito próximos), uma vez que a relação entre a internet, seus usuários e os dados, principalmente, é carro-chefe da multidisciplinaridade que pretendemos com áreas do saber voltadas à tecnologia da

³⁷ Web Semântica. W3C Brasil. 20[?]. Disponível em: <<https://www.w3c.br/Padroes/WebSemantica>>. Acesso em: 18 abr. 2022.

informação. Das tantas relações possíveis entre nossas áreas, uma que nos chama atenção neste trabalho são as estatísticas nas humanidades digitais, i.e., a existência de diferentes formas de visualização de informações.

Sobre essa coexistência entre as humanidades digitais e as estatísticas, Arnold e Tilton argumentam que “a formulação teórica de gráficos fornecidos pelas estatísticas oferece ideias conceituais com as quais as humanidades digitais podem lidar e manusear”³⁸, ao passo que, posteriormente, adicionam que “usar uma linguagem de programação cujo propósito geral seja estatístico aumenta a quantidade de aproximações metodológicas ao estudo de dados nas humanidades”³⁹ (ARNOLD & TILTON, 2019, Cap. 24, traduções nossas). Ao passo que as estatísticas nas humanidades são, inclusive por estes autores, diretamente ligadas ao uso de *Python* e *R*, o termo “*graphics*”, gráficos, também acaba se expandindo a outros usos das estatísticas e da raspagem e produção de dados, como os grafos.

A existência de quaisquer dessas formas de visualização de dados remete diretamente a padrões completamente diferentes de compreensão das fontes. Um gráfico, um mapa, um histograma e um diagrama oferecem cenários estatísticos diferentes um do outro, mesmo sobre um mesmo documento. Um grafo, igualmente, contém informações visuais que até sua implementação não seriam totalmente contempladas. A existência das estatísticas realiza essa substancialidade na interpretação de fontes e documentos históricos. É importante que, para cada uso seja entendido pelo/a intelectual suas próprias necessidades e almejos, uma vez que dificilmente conseguirá extrair a mesma informação a partir de modelos estatísticos diferentes, portanto, pretendemos expandir um pouco o debate sobre grafos. Ao falar sobre formas de visualizações nas humanidades digitais, Sinclair et al. expõe a ideia de que

[...] uma visualização que produz uma única saída para um certo corpo de materiais tem uma utilidade limitada; uma visualização que fornece muitas formas de interagir com os dados, visto de diferentes perspectivas, é melhor; um visualização que contribui para novas e emergentes formas de entender o material é a mais desejável. (SINCLAIR et al., 2013, documento *online*, tradução nossa)⁴⁰

De forma resumida, enquanto gráficos apresentam visualmente, através de colunas e linhas, majoritariamente valores numéricos e diagramas procuram produzir um “mapa conceitual” a partir de palavras-chave, a fim de resumir uma ideia ou uma apresentação, os

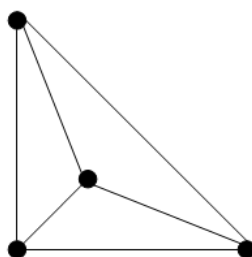
³⁸ Originalmente, “The theoretical formulation of graphics provided by statistics offers concepts ideas for the digital humanities to draw from and grapple with”

³⁹ Originalmente, “Using a general-purpose statistical programming language greatly increases the available set of methodological approaches to studying humanities data”

⁴⁰ Originalmente, “a visualization that produces a single output for a given body of material is of limited usefulness; a visualization that provides many ways to interact with the data, viewed from different perspectives, is better; a visualization that contributes to new and emergent ways of understanding the material is best.”

grafos partem de outro princípio. Grafos são estruturas matemáticas que expressam visualmente as conexões entre um conjunto de dados, contendo duas informações processuais principais: os nós (*nodes*), ou vértices, comumente visualizados como círculos no grafo, e as arestas (*edges*), geralmente visualizados como linhas ou setas que conectam os círculos. Teoricamente, se tivermos um conjunto de nós $N = \{1, 2, 3, 4, 5\}$ e queremos conectá-los de alguma forma, a partir de um grafo e uma relação comum, eles podem ser vinculados um a outro através das arestas e da relação $A = \{1, 2\}, \{1, 4\}, \{2, 4\}$, etc (vínculos a serem definidos de acordo com o/a autor/a do grafo). Um grafo simples funciona assim:

Figura 1. Grafo básico, com quatro nós e seis arestas.



Fonte: Wikipédia, 2005⁴¹.

Fica subentendido que, para uma análise de redes sociais, os grafos são a forma de visualização mais interessante a se utilizar. Digo isso porque, entendendo-os na sua forma e complexidade, percebo que as perspectivas presentes em um grafo ultrapassam aquelas presentes nos gráficos, mapas mentais e etc., uma vez que possibilitam interpretações mais profundas sobre as informações; e, especialmente quando vinculados às redes sociais, grafos possibilitam análise através de algoritmos para entender comportamento e gerar comunidades, característica que outras formas de visualização não possuem e são extremamente importantes à análise de redes sociais. Além disso, enquanto aquelas visualizações, por vezes, se prendem a uma limitação visual (i.e., são dados direta e especificamente informados pelo usuário na hora de criar um gráfico), os grafos são mais maleáveis, tendo diversas especificidades estatísticas e filtros para facilitar sua manipulação e organizar sua visualização de forma mais abrangente. Talvez essa seja a mais importante diferença entre essas diferentes conformações. Portanto, a partir de Sinclair et al., entendo os grafos como um dispositivo de criação de novas e emergentes maneiras de visualizar e interagir com a operação histórica e com a escrita da história, transformando-os em um produto cujo valor metodológico é muito maior.

⁴¹ Disponível em: <https://pt.wikipedia.org/wiki/Ficheiro:Grafo_k4_plano.PNG>. Acesso em: 18 abr. 2022.

Sobre a utilização de redes sociais virtuais, Regattieri et al. consideram que:

A relevância no contexto contemporâneo de sites de rede on-line serve como meio para interpretar as ações políticas e coletivas, é por isso que o Twitter é o nosso "campo" de trabalho. Consideramos que essa rede social um rico terreno de disputa, observando as muitas revoltas em todo o mundo: *#OccupyWallStreet*, *#15M*, *#OccupyGezy*, *#VemPraRua* e *#NãoVaiTerCopa*. (REGATTIERI et al. 2014, p. 2)

Essa consideração tem validação neste trabalho, pois consideramos que o Twitter enquanto rede social virtual pode ser um campo de disputa muito importante para a historiografia digital e para a produção de fontes nascidas digitais. *Hashtags*, interações entre perfis e criação de comunidades são objetos de análise importantes para elaborar ou minuciar investigações que elevam o ciberespaço à condição de fonte. Os grafos, para tanto, são o ponto onde esses objetos se tornam visuais, não apenas textuais ou numéricos; em *softwares* que os produzem, são o mapa que localiza, manipula e ajuda a interpretar essas comunidades e seus debates. Analisar e trabalhar com redes sociais virtuais é, para todos os fins, acompanhar o debate público ocorrido na internet de forma corrente e entender a atuação de diferentes grupos em um mesmo bloco de disputa de maneira muito mais quantitativa do que qualquer outra.

Dados apresentados pela própria plataforma informam que cerca de 217 milhões de usuários acessaram o *Twitter* diariamente no final de 2021⁴², número que, usualmente, seria impossível conseguirmos quantificar em uma pesquisa tradicional (aquelas onde as pessoas voluntariamente respondem certos questionamentos). Não quero dizer que podemos ter dados quantitativos e qualitativos provenientes de 217 milhões de pessoas, mas que, mesmo que seja o equivalente a 0,001% disso, já teríamos um conjunto muito significativo para fazer uma análise sobre uma temática. Assim, se tivermos em mãos a interação de 2 mil usuários diferentes sobre um assunto, temos, no mínimo, uma comunidade de duas mil opiniões simultâneas funcionando em nossa máquina e servindo como base de dados para nossas pesquisas. Serge Noiret (2014, p. 43), sem citar essas redes sociais virtuais, argumenta sobre como a *web* pode ser e é utilizada para coletar memórias e testemunhos de comunidades dispersas em diferentes localizações do globo, fazendo com que haja a possibilidade de comparação entre diferentes realidades, mesmo que não tenhamos contato *in loco* com elas. E, se tratando de redes virtuais e grafos, pode-se dizer que a ideia conceitual funciona de maneira bem semelhante.

⁴² Twitter tem crescimento de 13% em usuários enquanto Facebook fica estagnado. Tecnoblog. 2022. Disponível em: <<https://tecnoblog.net/noticias/2022/02/10/twitter-tem-crescimento-de-13-em-usuarios-enquanto-facebook-fica-e-stagnado>>. Acesso em: 18 de abr. 2022.

A correlação entre os dados de um grafo se propõe a entender e interpretar (e mesmo comparar) a opinião de X (X sendo uma variável numérica) usuários, mas também produzir resultados diferentes para esses dados e capacitá-los como ferramentas que evidenciam não somente *o que pensa* uma comunidade, mas como elas se comportam, quais são suas características definidoras enquanto comunidades virtuais distintas, bem como análises extra-textuais. Por exemplo, anteriormente neste trabalho, citamos a existência da ferramenta de análise de sentimento das redes, uma amostra do pode ser feito a partir de dados provenientes de usuários nas redes sociais. Essas informações são, portanto, camadas que dão origem às perspectivas. Israel Aquino, em curto artigo sobre a análise de redes sociais, argumenta que

Para além da elaboração das matrizes e a geração dos respectivos grafos, outra contribuição fundamental deste método é a produção dos coeficientes ou estatísticas de análise, denominadas métricas, que permitem evidenciar algumas das características dessas redes, colocando em evidência alguns tipos de relação que se estabelecem, ou destacando a atuação de um determinado agente no grupo a partir da posição por este assumida. (AQUINO, 2019?, documento *online*)⁴³

Para gerar um grafo, primeiro é importante ter os dados a serem conectados bem definidos, postos em uma tabela com algumas informações específicas pré-estabelecidas⁴⁴. Após isto, existem alguns *softwares*, *plugins* ou até mesmo sites para auxiliar na criação destes grafos. Há o NodeXL, um pacote de *plugin* disponível para o *Microsoft Excel*, o editor de planilhas que a maioria das pessoas já conhece; além disso, há um site, muito mais básico e mais manual (mas que também requer um certo conhecimento sobre alguns termos e conceitos inerentes à teoria dos grafos) chamado *Graph Online*⁴⁵; há, enfim, o *software* mais utilizado, inclusive neste trabalho, para a produção de grafos, chamado *Gephi*, feito em código aberto e gratuito⁴⁶, que permite a criação desses grafos a partir de arquivos exportados pelo usuário para dentro do programa. A partir disso, o usuário é livre para trabalhar da forma que quiser com os dados. Há caminhos mais populares para os quais pesquisadores/as das ciências humanas devem se dirigir (i.e., estatísticas e filtros que são mais úteis para nossa

⁴³ AQUINO, Israel. "Análise de redes sociais". In: CLIOMATICA - Portal de História Digital e Pesquisa. Disponível em: <http://lhs.unb.br/cliomatica/index.php/An%C3%A1lise_de_redes_sociais>. Acesso em: 18 abr. 2022.

⁴⁴ Para uma introdução sobre o uso do Gephi, bem como um tutorial sobre instalação e uso *software*, a leitura que mais indicamos (disponível em inglês) é a de Martin Grandjean, um expoente importante do estudo dos grafos nas humanidades digitais, cf. GRANDJEAN, M. GEPHI – Introduction to Network Analysis and Visualization. Martin Grandjean. 2015. Disponível em: <<http://www.martingrandjean.ch/gephi-introduction/>>. Acesso em: 18 abr. 2022.

⁴⁵ cf. Graph Online. Disponível em: <<https://graphonline.ru/pt/>>. Acesso em: 18 abr. 2022.

⁴⁶ O site do *software* é <<https://gephi.org/>>, mas recomendamos a leitura do manual de instalação e utilização feito por Martin Grandjean, inclusive como auxiliar.

interpretação), mas o Gephi concede liberdade o suficiente para o/a usuário/a trabalhar como quiser com os conjuntos.

A seguir, demonstraremos a trajetória percorrida para produção de um grafo, desde a concepção dos dados a partir de uma raspagem no *Twitter* feita via *script* de *Python*, passando pela tabela que é constituída com esses dados e, enfim, como eles se comportam enquanto grafos, dentro do Gephi. A ideia é demonstrar a prática produtora de um grafo e, posteriormente, projetos bem estabelecidos cujos grafos são o dispositivo central do produto final, que utilizem-os como fonte de estudo e objeto de análise.

4.1. Como produzir grafos

Depois de feitas as instalações de todas as ferramentas que serão necessárias para o/a usuário/a iniciar esta odisséia (a citar: instalação e configuração inicial do *Python*, do *Twint* e do *Gephi*, no mínimo), a primeira etapa é a raspagem de dados (de *tweets*), feita em *Python*. Para isso, cria-se um *script* com as necessidades que o/a autor/a desejar. O *Twint*, por questões internas do *Twitter*, tem uma certa limitação por *script* de raspagem, não conseguindo fazer uma coleta de dados com muitos dias (pelo menos não nos testes preliminares feitos para este trabalho). Neste exemplo, o *script* (salvo como *scriptAlessandro.py*) terá início dia 14/07/2021 e fim 19/07/2021, o termo a ser pesquisado é “@Sen_Alessandro CPI” e o resultado será salvo em uma tabela .csv (arquivo denominado *senAlessandro.csv*), que pode ser editada via *Microsoft Excel* ou *Google Sheets*. A ideia é coletar publicações que marcaram o usuário @Sen_Alessandro, o senador da república Alessandro Vieira e citaram o termo CPI (contextualmente, a CPI da COVID que acontecia naquele momento). Na captura de tela, o *script* está pronto e cada linha do código é explicada após a #cerquilha (toda linha escrita em *Python* não é “lida” na sua execução após a cerquilha). O *software* utilizado para escrevê-lo é o *Notepad++*⁴⁷, uma versão “atualizada” do clássico bloco de notas do computador que é comumente utilizado para programação.

⁴⁷ Gratuito para download em <<https://notepad-plus-plus.org/downloads/>>. Acesso em: 19 de abr. 2022.

Figura 2. Script do Twint em Python.

```

scriptAlessandro.py
1 import twint
2
3
4 c = twint.Config()
5 c.Search = "@Sen_Alessandro CPI" #Termo a ser pesquisado
6 c.Until = "2021-07-19" #Data final a ser pesquisada
7 c.Since = "2021-07-14" #Data inicial a ser pesquisada
8 c.Store_csv = True #Criar um documento csv
9 c.Output = "senalessandro.csv" #Nome do documento a ser criado
10
11 twint.run.Search(c)
    
```

Fonte: Figura do autor. Captura de tela.

A seguir, a execução é feita via *Prompt de Comando*⁴⁸ do computador usando o gatilho “python scriptAlessandro.py” e, na mesma pasta onde o arquivo .py está salvo, será criado um arquivo de tabela .csv. Neste caso, abrimos o arquivo via *Google Sheets* e, nele, há inúmeras informações, muitas delas inúteis para este processo. Fazendo uma breve limpeza na tabela, deixamos as duas colunas que interessam para criar o arquivo que irá para o Gephi e mais algumas que serão úteis enquanto metadados (como contagem de curtidas, *retweets* e respostas a cada *tweet*, bem como o *link* para acessar publicação na própria rede social).

Figura 3. Tabela “antes”, com todos os dados que são coletados pelo script.

A1	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	id	conversation_id	created_at	date	time	timezone	user_id	User	name	place	Message	language	mentions	urls	photos
2	1417269262729	1417258281878	2021-07-19 20:4	2021-07-19	20:44:58	-300	1144252684540	passivo7747596	o genro que sua mãe pediu	@manueloma pt					
3	1417268012550	1417258281878	2021-07-19 20:4	2021-07-19	20:40:00	-300	9694450118793	emanuelomanu	manu	@Passivo77475 pt					
4	1417265471380	1417258281878	2021-07-19 20:2	2021-07-19	20:29:54	-300	1144252684540	passivo7747596	o genro que sua mãe pediu	@matheusantant pt					
5	1417265122825	1417258281878	2021-07-19 20:2	2021-07-19	20:28:31	-300	1144252684540	passivo7747596	o genro que sua mãe pediu	@manueloma pt					
6	1417265052370	1417258281878	2021-07-19 20:2	2021-07-19	20:28:14	-300	33375814	matheusantana	Matheus Santana #ForaBolsona	@Passivo77475 pt					
7	1417260052562	1417258281878	2021-07-19 20:0	2021-07-19	20:08:22	-300	9694450118793	emanuelomanu	manu	@Passivo77475 pt					
8	1417259510746	1417258281878	2021-07-19 20:0	2021-07-19	20:06:13	-300	1144252684540	passivo7747596	o genro que sua mãe pediu	@camarotodacp pt					
9	1417228613397	1417228613397	2021-07-19 18:0	2021-07-19	18:03:26	-300	1157882808641	hp55804711hp	👉👉👉👉👉👉	# Parece que o pt					
10	1417226798157	1417198984771	2021-07-19 17:5	2021-07-19	17:56:14	-300	98220950	marcelleims	Marcelle Mello	@Lili_afns @Se pt					
11	1417217875128	1417198984771	2021-07-19 17:2	2021-07-19	17:20:46	-300	7982017435322	dani_sant30	Daniela Santos	@Lili_afns @Se pt					
12	1417207474071	1417160508424	2021-07-19 16:3	2021-07-19	16:39:26	-300	1227608659955	abade_andre	Andre Abade	@Ange156064; pt					
13	1417204132042	1417097433084	2021-07-19 16:2	2021-07-19	16:26:10	-300	169866262	scr19872008	NAURICIO 1445	@OmarAzizSen pt					
14	1417200044776	1416786720004	2021-07-19 16:0	2021-07-19	16:09:55	-300	1004893740274	nadiacouloubern	Nádia Couto Bem	@VemPraRua_1 pt					
15	1417197829789	1417160508424	2021-07-19 16:0	2021-07-19	16:01:07	-300	1004893740274	nadiacouloubern	Nádia Couto Bem	@Sen_Alessanc pt					
16	1417197447210	1417160508424	2021-07-19 15:5	2021-07-19	15:59:36	-300	1344054021703	montezuma_sak	francisco montezuma sales	@OmarAzizSen pt					
17	1417196377457	1417160508424	2021-07-19 15:5	2021-07-19	15:55:21	-300	1344054021703	montezuma_sak	francisco montezuma sales	@frankseixas @ pt					
18	1417191989926	1417161875977	2021-07-19 15:3	2021-07-19	15:37:55	-300	1375418984170	11982soares	Felipe SF	@cynaramenezzi pt					
19	1417188027823	1417188027823	2021-07-19 15:2	2021-07-19	15:22:10	-300	1379554879161	fonscacadaniel01	Daniel Fonseca	@CarlaBigatto fi pt					
20	1417185373996	1417185373996	2021-07-19 15:1	2021-07-19	15:11:37	-300	33642774	lejoq	Ministro do Gov	@ethonlex					
21	1417174341314	1417174341314	2021-07-19 14:2	2021-07-19	14:27:47	-300	1224105634519	rosianemariar1	Rosiane Maria Araujo de Oliveira	@randolfcap @ pt					
22	1417172895411	1417160508424	2021-07-19 14:2	2021-07-19	14:22:02	-300	1227608659955	abade_andre	Andre Abade	@Tonifaria2 @f pt					
23	1417169676023	1417160508424	2021-07-19 14:0	2021-07-19	14:09:15	-300	1362287374168	victorcabata	Victor C.	@Tonifaria2 @f pt					
24	1417167084539	1417160508424	2021-07-19 13:5	2021-07-19	13:58:57	-300	1362287374168	victorcabata	Victor C.	@Tonifaria2 @f pt					
25	1417165219969	1417160508424	2021-07-19 13:5	2021-07-19	13:51:32	-300	1413636899139	tonifaria2	Sugisumundo da Depressão	@Beguinho9 @ pt					
26	1417164848455	1417160508424	2021-07-19 13:5	2021-07-19	13:50:04	-300	1349446797148	beguinho9	Beguinho13	@Tonifaria2 @: pt					
27	1417164598928	1417160508424	2021-07-19 13:4	2021-07-19	13:49:03	-300	1383735628618	carlos08704643	Carlos	@Sen_Alessanc pt					
28	1417162866113	1417160508424	2021-07-19 13:4	2021-07-19	13:42:11	-300	1413636899139	tonifaria2	Sugisumundo da Depressão	@Sen_Alessanc pt					
29	1417161152929	1417160508424	2021-07-19 13:3	2021-07-19	13:35:23	-300	1156731133549	veralviasiquei1	VERA LÍVIA SIQUEIRA	@Sen_Alessanc pt					
30	1417158976152	1417158976152	2021-07-19 13:2	2021-07-19	13:26:44	-300	1259605100194	francis47333003	Framor	A CPI precisa te pt					
31	1417157338188	141706506103	2021-07-19 13:2	2021-07-19	13:20:13	-300	73160697	jozileury	Jojo, uma Feérica do Cunó Utinç	@marlusromero pt					
32	1417156693985	1415095076880	2021-07-19 13:1	2021-07-19	13:17:40	-300	134558324209	danieldantasdr2	Daniel Dantas Dória	@SimoneTebet pt					
33	1417154427362	1417097433084	2021-07-19 13:0	2021-07-19	13:08:39	-300	1373344476429	nenaz8914013	Nena # FORA BOLSON	@OmarAzizSen pt					
34	1417138352026	1417134186901	2021-07-19 12:0	2021-07-19	12:04:46	-300	12426136111061	fabricio21089024	Fabricio #ForaBolsona	@diegoescostec pt					
35	1417130530437	1417130530437	2021-07-19 11:3	2021-07-19	11:33:42	-300	1044774362426	ruissantos	Rui Sérgio	@OmarAzizSen pt					

Fonte: Figura do autor. Captura de tela.

⁴⁸ Para saber mais sobre o que é o Prompt de Comando de um computador, cf. O que é o Prompt de Comando do Windows?. Oficina da Net. Disponível em: <<https://www.oficinadonet.com.br/windows10/28262-o-que-e-o-prompt-de-comando-do-windows>>. Acesso em: 19 abr 2022.

Figura 4. Tabela “depois”, com colunas dispensáveis ocultadas.

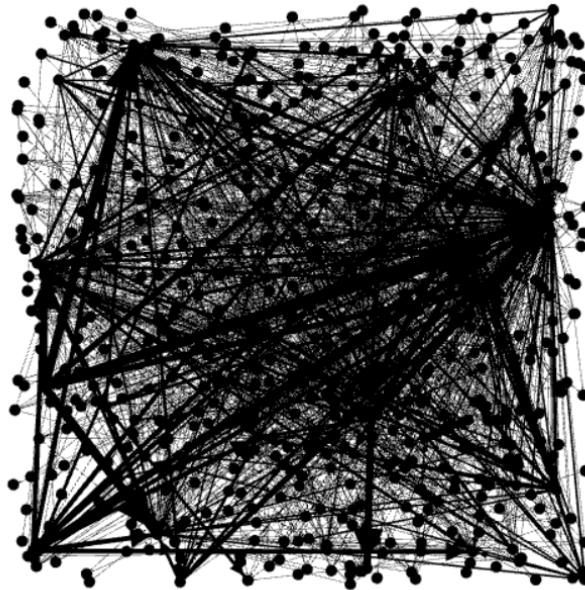
	A	C	H	I	J	K	M
1	User	Message	replies_count	retweets_count	likes_count	hashtags	link
2	passivo77475968	@emanueloumanu @camarotedacpi @Sen_Alessand	0	0	0		https://twitter.com
3	emanueloumanu	@Passivo77475968 @camarotedacpi @Sen_Alessan	1	0	0		https://twitter.com
4	passivo77475968	@matheusantana @camarotedacpi @Sen_Alessandri	0	0	2		https://twitter.com
5	passivo77475968	@emanueloumanu @camarotedacpi @Sen_Alessand	1	0	0		https://twitter.com
6	matheusantana	@Passivo77475968 @camarotedacpi @Sen_Alessan	1	0	6		https://twitter.com
7	emanueloumanu	@Passivo77475968 @camarotedacpi @Sen_Alessan	1	0	1		https://twitter.com
8	passivo77475968	@camarotedacpi Creio q o @Sen_Alessandro investic	2	0	8		https://twitter.com
9	hp55804711hp	👉 Parece que o Brasil pode voltar a sonhar! 🇺🇦 Será	0	0	2	[cpidapandemia	https://twitter.com
10	marcellems	@Lili_afns @Sen_Alessandro Saudades, CPII	1	0	3		https://twitter.com
11	dani_sant30	@Lili_afns @Sen_Alessandro Melhores momentos da	1	0	1		https://twitter.com
12	abade_andre	@Angel1500642247 @ToniFaria2 @Sen_Alessandro	0	0	0		https://twitter.com
13	scr19872008	@OmarAzizSenador Aqui da cidade de Sirinhaém/PE,	0	0	0		https://twitter.com
14	nadiacoutobem	@VemPraRua_br @Sen_Alessandro Cadê a moral de	0	0	0		https://twitter.com
15	nadiacoutobem	@Sen_Alessandro Senador, o senhor foi a PIOR dece	0	0	0		https://twitter.com
16	montezuma_sales	@Gmaraujo3 @Sen_Alessandro @EDSONVINHEDO	0	0	0		https://twitter.com
17	montezuma_sales	@frankeixas @Sen_Alessandro Acompanhamento a CPI d	0	0	2		https://twitter.com
18	f1992soares	@cynaramenezes Isso é material BOMBA para a CPI,	0	0	0		https://twitter.com
19	fonsecadanielo1	@CarlaBigatto fala pro @renancalheiros avisar o @ra	0	0	0		https://twitter.com
20	lexjq	Ministro do Gov federal está no @flowpdc dizendo q o	0	0	1		https://twitter.com
21	rosianemariaar1	@randolfeap @OmarAzizSenador @renancalheiros @	0	0	0		https://twitter.com
22	abade_andre	@ToniFaria2 @Sen_Alessandro Acho que o @ToniFa	2	0	1		https://twitter.com
23	victorbatista	@ToniFaria2 @Beginho9 @Sen_Alessandro Não est	1	0	0		https://twitter.com
24	victorbatista	@ToniFaria2 @Beginho9 @Sen_Alessandro ANTIC	1	0	1		https://twitter.com
25	tonifaria2	@Beginho9 @Sen_Alessandro Nao lvo a pergunta?l	2	0	0		https://twitter.com
26	beginho9	@ToniFaria2 @Sen_Alessandro Ja ouviu falar em dep	1	0	19		https://twitter.com
27	carlos08704643	@Sen_Alessandro Estou satisfeito com sua postura n	0	0	5		https://twitter.com
28	tonifaria2	@Sen_Alessandro E na CPI senador, vai ouvir os ges	6	0	12		https://twitter.com
29	veralviasquei1	@Sen_Alessandro Não apoio nada que venha do Sen	4	0	5		https://twitter.com
30	francis47333003	A CPI precisa ter a coragem de ser o que é, SEN. da l	0	0	2		https://twitter.com
31	jozifeury	@marlusromero @Cissafb @guilherme_amado @prc	1	0	4		https://twitter.com
32	danieldantasdr2	@SimoneTebetms @Sen_Alessandro @tabataamaral	0	0	1		https://twitter.com
33	mena28914013	@OmarAzizSenador CPI SENDO PRESIDIDA POR U	1	0	1		https://twitter.com
34	fabrcio21089024	@diegoescosteguy Interessante e da pra comparar os	0	1	1		https://twitter.com
35	ruissantos	@OmarAzizSenador @senadorhumberto @Sen Ale	0	0	0		https://twitter.com

Fonte: Figura do autor. Captura de tela.

Destas, as colunas “User” e “Message”, que validam o usuário que publicou a mensagem e qual mensagem foi publicada, são importantes. Com essas colunas tendo esses títulos, podemos automatizar a transformação dessa tabela em um arquivo para o Gephi pela ferramenta⁴⁹ criada por Luca Hammer, um programador que trabalha com criação de grafos. Nesta, quando um arquivo .csv tem as colunas nomeadas acima, ele é automaticamente transformado para um arquivo .gdf, de leitura do Gephi. Com o arquivo .gdf salvo no computador, basta abri-lo no Gephi e teremos o início de um grafo, na sua forma bruta (ou seja, é um grafo sem métricas, estatísticas ou filtros, apenas um grande quadro cheio de círculos e linhas pretos).

⁴⁹ A ferramenta *CSV to GDF converter* está disponível no site <<https://csvtogdf.glitch.me/>>, sendo este de código aberto, capacitando outras pessoas a “copiarem” o código e publicarem-no em seus próprios domínios. Eu, enquanto usuário principiante, prefiro, por ora, utilizá-la no domínio disponibilizado por Luca Hammer.

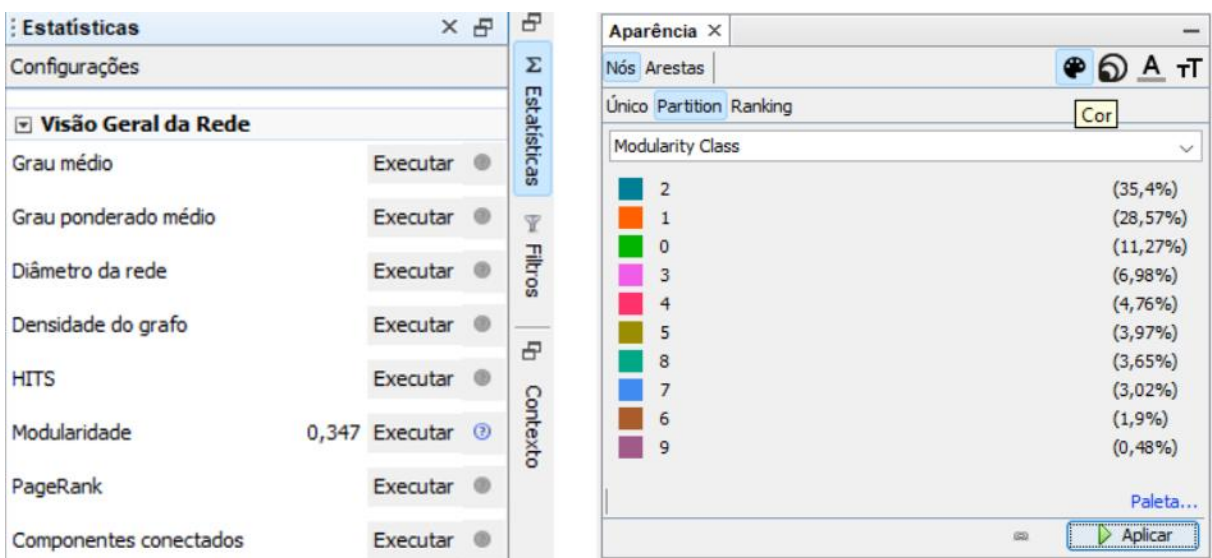
Figura 5. Grafo visto no Gephi, sem tratamento.



Fonte: Figura do autor. Captura de tela.

Sobre as métricas importantes a serem consideradas, no menu “Estatísticas”, à direita no próprio Gephi, é importante configurar a modularidade do grafo, um algoritmo para detecção de *comunidades*. O grau de modularidade permitirá ao/à usuário/a criar, a partir desse filtro, comunidades com cores distintas (facilitando a interpretação de quais são os usuários que, interagindo entre si, formam grupos particulares).

Figura 6. À esquerda, o menu de estatísticas e, à direita, a modularidade aplicada às cores dos nós.

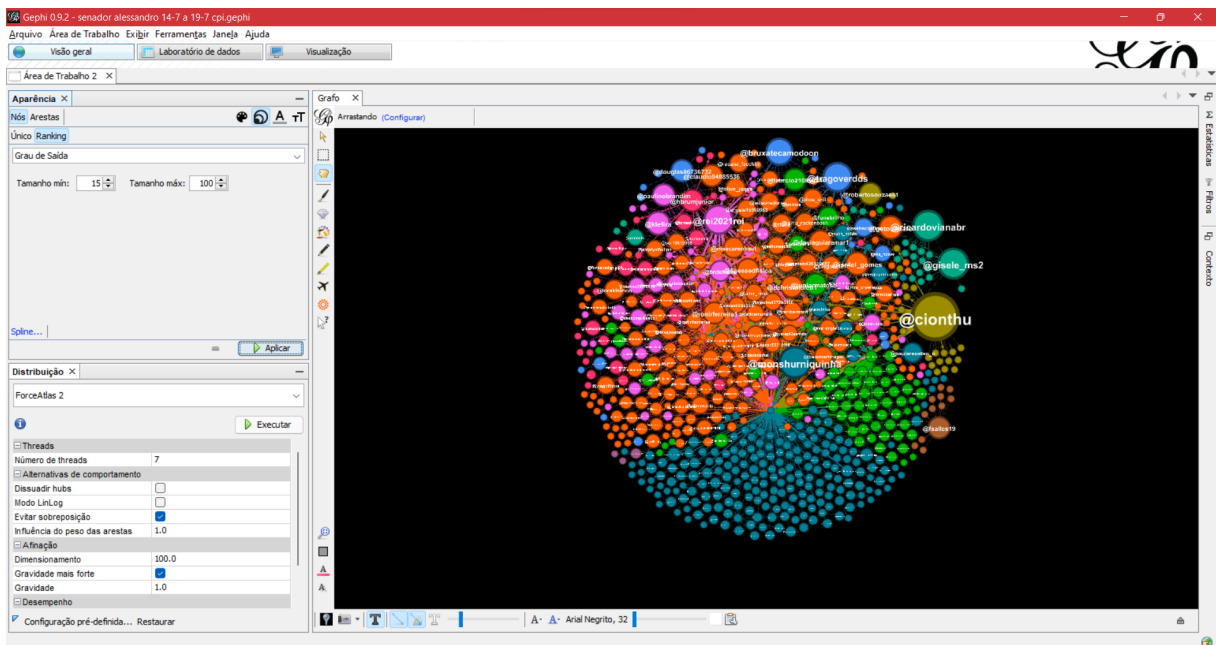


Fonte: Figura do autor. Captura de tela.

Pode-se perceber, a partir desta figura, que ao lado das cores dispostas pelo Gephi para cada comunidade, há a porcentagem de nós que terão essa cor. Muitas vezes, essas altas porcentagens podem ser vinculadas aos grupos periféricos, que não foram captados por nenhuma comunidade específica e não interagiram (ou sofreram interações) o suficiente para serem categorizados pelo algoritmo. Após isso, deve-se escolher um modo de distribuição (menu que está localizado abaixo do menu de aparência, no Gephi), sugerindo-se *ForceAtlas 2* ou *Fruchterman Reingold*⁵⁰ e, então, já haverá um resultado preliminar do grafo. Após isso, o grafo pode ter o tamanho dos nós alterado, no menu aparência, utilizando alguma métrica que o/a autor/a desejar. Para este exemplo, escolhemos filtrar o tamanho dos nós a partir do grau de saída (para entendermos de onde estão saindo as interações, i.e., quais são os usuários que mais têm interações dentro de cada comunidade), com tamanho mínimo 15 e tamanho máximo 100. Outra métrica aplicada foi a exibição textual de cada nó (com tamanho respectivo), para tornar mais fácil saber quem seu/sua autor/a.

O resultado geral, aplicando essas métricas e filtros, fica desta forma no *software*:

Figura 7. Visão geral do grafo no Gephi após tratamento.



Fonte: Figura do autor. Captura de tela.

⁵⁰ Ambos são nomes dados aos algoritmos que fazem a distribuição dos nós e arestas do grafo para torná-los mais visualizáveis, há outros, mas, em geral, recomenda-se utilizar estes dois por serem os mais organizados quanto à forma final que o grafo tomará.

Após o grafo “pronto”, ele pode ser salvo como uma imagem na aba “Visualização” e, na aba “Laboratório de dados”, ainda há a possibilidade de perceber alguns metadados do próprio Gephi. A partir daqui, os usos dos dados, metadados, nós e arestas do grafo cabe ao/à pesquisador/a definir, mas uma utilização conjunta do grafo, das métricas que ele oferece e da tabela .csv (com os *links* de acesso ao *tweet*, a contagem de *retweets* e curtidas) leva a um resultado muito interessante.

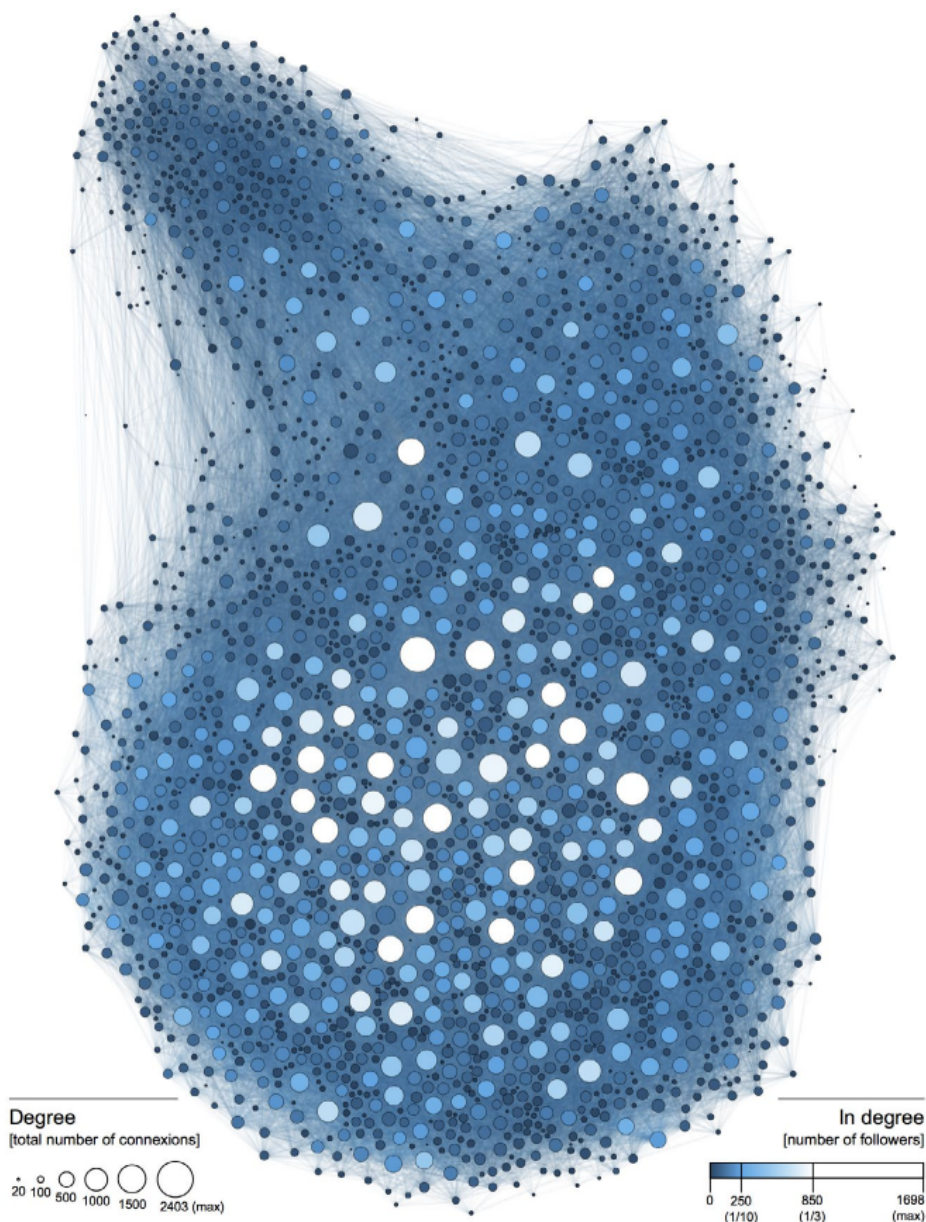
4.2. Usos dos grafos na historiografia

Sobre pesquisas que envolvem o uso de grafos, já foram citados neste trabalho dois pesquisadores, Martin Grandjean e Luca Hammer, bem como os projetos do LABIC. Há uma feliz coincidência entre esses três: enquanto Grandjean já trabalhou com criação de redes de relacionamento a partir de uma *#hashtag*, Luca Hammer fez uma extensa pesquisa sobre perfis de pessoas verificadas pela plataforma e, o LABIC possui projetos que envolvem termos e palavras-chave para a produção dos grafos. Cada um com uma forma diferente de manipular os dados, constroem projetos, narrativas e interpretações específicas e muito interessantes.

Grandjean, no seu site pessoal, voltado à pesquisa em *digital humanities*, além de artigos didáticos sobre o uso do Gephi, também produziu uma pesquisa, em 2016, mapeando a comunidade de humanidades digitais no Twitter. No seu trabalho (p. 2)⁵¹, ele explica que foram mapeados 2,500 perfis que se encaixam nessa “comunidade de práticas” e possuem uma mínima relação entre si (i.e., uma corrente de perfis que se seguem ou são seguidos). Para visualização, o autor produz um grafo que centraliza os perfis com maior número de interações (número de conexões e de seguidores).

⁵¹ Além do texto na íntegra, outras informações sobre este trabalho podem ser encontradas em: GRANDJEAN, M. [Network analysis] Mapping the Digital Humanities Community on Twitter. 2016. Disponível em: <<http://www.martingrandjean.ch/network-analysis-mapping-digital-humanities-community-twitter/>>. Acesso em: 19 abr. 22.

Figura 8. Grafo feito por Martin Grandjean.



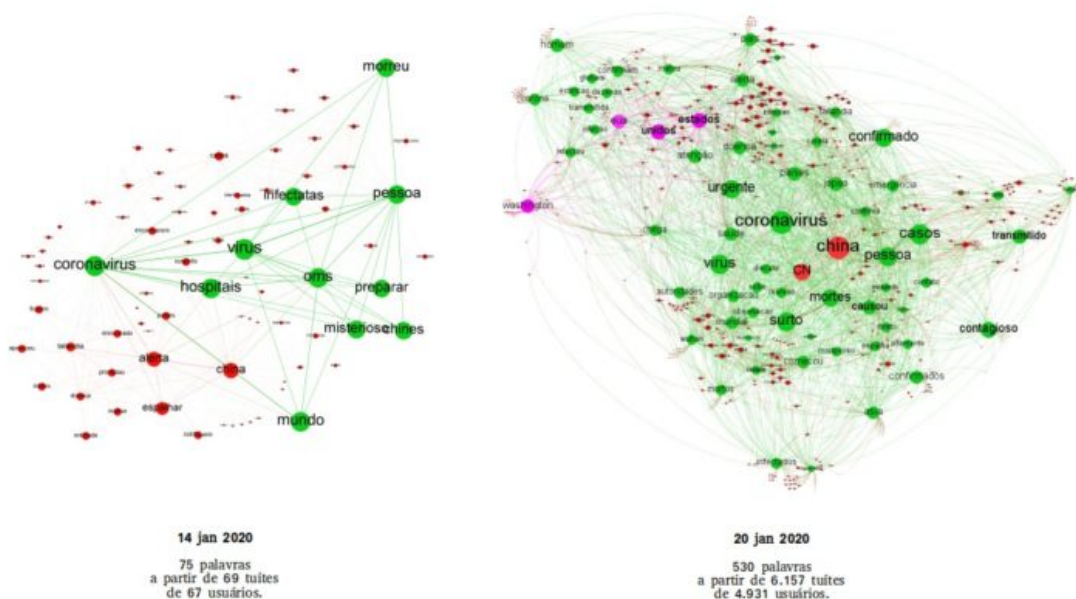
Fonte: GRANDJEAN, M. A social network analysis of Twitter: Mapping the digital humanities community. *Cogent Arts & Humanities*, v. 3, 2016.

Percebe-se como o autor utiliza duas métricas diferentes para definir e demonstrar duas estatísticas específicas. O produto final é, portanto, além de complexo, bastante didático. Após a leitura do seu artigo, visualizar as informações dessa forma torna o texto muito mais fácil de ser entendido e as informações que Grandjean cita, melhor observáveis. O grafo

cumpra uma função de facilitação para o/a leitor/a entender as informações exploradas pelo/a autor/a.

Fabio Malini, do LABIC, em 2020 produziu um trabalho sobre as narrativas eclodidas no *Twitter* sobre o covid-19. Além de uma produção textual sobre como funciona essa guerra de narrativa nas redes sociais virtuais, Malini utilizou de alguns grafos para explicitar melhor seus argumentos. No seu trabalho⁵², o autor manuseia alguns grafos que utilizam palavras-chave e outros partindo de usuários e, somando-se a estes, também utiliza de gráficos. O fato imprescindível de ser notado é que, mais do que um grafo que expõe as informações estabelecidas no corpo textual, o autor também demonstra como utilizá-los para fazer acompanhamento de certas demandas, ao mostrar, por exemplo, grafos comparativos entre como a rede se comportou sobre o mesmo assunto em dois dias diferentes:

Figura 9. Gráficos que mostram a *timeline* das redes de narrativas nos dias 14 e 20 de janeiro de 2020.



Fonte: MALINI, F. Quando tudo parecia ser tão distante daqui: a eclosão das narrativas sobre covid-19. **LABIC - Laboratório de Estudos sobre Imagem e Cibercultura**. 2020. Disponível em:

<<https://www.labic.net/cartografia/quando-tudo-parecia-ser-distante-daqui/>>. Acesso em: 19 de abr. 2020.

A partir desses dois exemplos de grafos, que carregam as características trazidas por este trabalho, é perceptível como a utilização desses dispositivos visuais pode servir como auxiliar e catalisador em pesquisas e projetos das ciências humanas e das humanidades

⁵² Para visualizar o artigo completo, com outros grafos e gráficos complementando o trabalho do autor, cf. MALINI, F. Quando tudo parecia ser tão distante daqui: a eclosão das narrativas sobre covid-19. LABIC - Laboratório de Estudos sobre Imagem e Cibercultura. 2020. Disponível em: <<https://www.labic.net/cartografia/quando-tudo-parecia-ser-distante-daqui/>>. Acesso em: 19 de abr. 2020.

digitais como um todo. Bastian et al. sugerem que, com o uso dos grafos gerados pelo Gephi, essas “visualizações são úteis para potencializar as habilidades humanas de percepção e encontrar características nas redes de estruturas e dados”⁵³, e, mais do que isso, são úteis para guiar os/as *experts*, autores/as e pesquisadores/as, na exploração de vastas e complexas redes de relacionamento (BASTIAN et al. 2009, p. 1).

Entretanto, avançando um pouco além da facilitação que os grafos sustentam nos trabalhos em que são utilizados, sua utilização a partir desses dados provenientes exclusivamente de redes sociais virtuais e, portanto, da internet, também avançam no debate sobre as fontes nascidas digitais (*born-digital archives*, ou FNDs), uma vez que, por si só, um grafo e todos os metadados que ele possui é potencialmente uma fonte (primária ou não). O debate sobre FNDs é muito bem estabelecido por Fábio Chang de Almeida. O autor diz, no início de seu texto, em 2011, que

Para os historiadores que buscam compreender o presente, negligenciar as fontes digitais e a Internet significa fechar os olhos para todo um novo conjunto de práticas, de atitudes, de modos de pensamento e de valores que vêm se desenvolvendo juntamente com o crescimento e popularização da rede mundial de computadores. (ALMEIDA, 2011, p. 12)

Os documentos que o autor chama de digitais exclusivos são de ordem imperativa à existência da internet, tratando-se de uma “enorme quantidade de informação que está sendo produzida e disponibilizada unicamente em formato digital” (idem, p. 19). Esses documentos, outrora, foram muito constituídos através de *blogs pessoais*, bem como a partir de plataformas de redes sociais como o *YouTube*, o *Facebook* e, obviamente, o *Twitter*. A efemeridade de certos conteúdos que existem exclusivamente na internet, claro, é consequência importante a ser lembrada quando pensamos nas FNDs, principalmente com conteúdos que não sentimos a necessidade de “salvar” (ou fazendo o *download*, ou salvando a página da web em um formato que ela possa ser vista *offline*, ou mesmo captando e passando o seu conteúdo para a nuvem). A realidade é que todas essas plataformas e sites são potencialmente produtoras de fontes primárias a partir de seus *born digital archives*. Um projeto de pesquisa pode ser iniciado tendo em vista exclusivamente materiais que só existam na internet, sem a necessidade de haver uma fonte primária física. Dessa forma, uma tabela de comentários de vídeos disponíveis no *YouTube* pode ser uma fonte primária tanto quanto uma tabela de dados extraídos do *Twitter*. Acrescentamos as palavras de Kirschenbaum ao dizer que “textos eletrônicos, arquivos, *feeds* e transmissões de todo tipo também são agora, indiscutivelmente,

⁵³ Originalmente, “Visualizations are useful to leverage the perceptual abilities of humans to find features in network structure and data.”

fontes primárias (como prova, não precisa-se procurar além das hashtags do Twitter como #Egypt ou #Obama).” (KIRSCHENBAUM, 2013, p. 4, tradução nossa⁵⁴).

Assim, a questão mais notável a ser levantada é a de que grafos são, para todos os fins, passíveis de serem utilizados como fontes, primárias ou secundários, em quaisquer trabalhos que sejam possíveis e queiram utilizá-los. A abundância de dados (COHEN & ROSENZWEIG, 2005; ALMEIDA, 2011; FICKERS, 2012) pode, inclusive, ser melhor interpretada e manipulada a partir desses dispositivos que aglomeram grandes coleções de dados em um pequeno espaço visual. As FNDs merecem receber uma reflexão sistemática dos/as acadêmicos/as das ciências humanas, podendo ser elas pedestais de futuros projetos desses/as intelectuais.

Além disso, fica proeminente pensar que a teoria da história e a escrita da história são diretamente afetadas (no bom sentido) pela utilização de FNDs, podendo sofrer alterações de metodologia e epistemologia suficientes para que possamos, eventualmente, entender a história digital como parte integrante definitiva da operação historiográfica. A história digital é, portanto, um braço forte imprescindível da disciplina histórica, devendo ser utilizado conjuntamente com outras áreas do saber do nosso campo, auxiliando na construção e percepção dos novos trabalhos. Urge, para tanto, a plena participação dos/as profissionais da história no espaço público deixado pela internet e pelas redes sociais virtuais, sendo fundamental que tenhamos colaboração e cooperatividade com outras áreas do saber para, junto da nossa própria cientificidade, consigamos desenvolver iniciativas úteis à sociedade e ajudar a construir um espaço virtual mais seguro e confiável. Tal integração dos/as historiadores/as nos espaços digitais também é fundamental para que tenhamos quórum suficiente de profissionais para (tentar) dar conta de todas as demandas que e esses locais têm, pois, como escreve Noiret, “o historiador encontra-se apto a levantar novas questões epistemológicas na análise do passado, depois do advento dos meios digitais. Todavia, apenas uma minoria de “historiadores digitais” domina os instrumentos que respondem às novas interrogações científicas” (2015, p. 33).

Enfim, é substancial perceber a teia que compõe o relacionamento entre a história digital, as humanidades digitais, as redes sociais virtuais e todos os instrumentos capazes de fornecer facilitações teórico-metodológicas para os/as pesquisadores/as e gerar (ou se aproveitar das) fontes nascidas digitais. A forma final dessa teia é a multidisciplinaridade e a

⁵⁴ Originalmente, “Electronic texts, files, feeds, and transmissions of all sorts are also now, indisputably, primary records (for proof one need look no further than Twitter hashtags such as #Egypt or #Obama).”

interconexão entre nós e diversas outras áreas e saberes específicos. Devemos, sempre, aproveitar essas oportunidades.

5. CONSIDERAÇÕES FINAIS

Não se pode esperar que uma só pessoa entenda um espectro tão grande de disciplinas reunidas nas humanidades digitais na profundidade requerida para desenvolver os métodos e entendimentos que são a promessa desse campo. (ARNOLD & TILTON, 2019, documento online, tradução nossa)⁵⁵

Pontualmente, é seguro dizer que o entendimento de história digital que temos hoje não é o mesmo que havia em 2005, a partir de Roy Rosenzweig e Daniel J. Cohen. A presunção sobre alguns aspectos que consternou os estudiosos daquele momento sofreu alterações, bem como, em que pese algumas preocupações, foram ultrapassadas. Conceitos sobre o campo de estudos e seus afazeres metodológicos, os/as profissionais que poderiam se interessar por ele, bem como a construção de pensamento que estes/as devem seguir estão em franca expansão. Hoje, a HD é um largo espaço, abrigando diversos debates e produzindo outros ainda, como o presente neste trabalho, a fim de mantermos a disciplina histórica e a pesquisa acadêmica atualizadas de quaisquer novidades tecnológicas que possam nos alcançar. Sem desmerecer - e inclusive incluindo - as discussões levantadas desde o momento que a HD se tornou uma realidade, o caminho percorrido até o presente é curto, mas promissor e, decerto, duradouro.

Embora, para muitos expoentes da HD, ela seja uma longa promessa, penso, a partir deste trabalho, que existem outras formas de percebê-la que não tão somente aquelas que recolhem-na a uma esfera de produtividade limitada, que terá fim ‘se não alcançar certa notoriedade enquanto é tempo’, ou que servirá simplesmente a um propósito - o de lidar com novidades até que elas não sejam mais tão novas assim. Produzo o raciocínio que espero da HD a possibilidade de, conjuntamente ao que chamamos de esfera tradicional da história, metodologicamente alterar certas concepções do que é escrever história e como se faz a operação historiográfica que De Certeau cita na sua obra.

Ademais, compreendo que, além de repassar o que há de novo no universo da computação e da modernidade tecnológica, cabe à HD ser a figura representante dessa Era Digital da historiografia, podendo, ela própria, fazer história enquanto campo de estudo independente, sem estar categorizado como um ‘braço’ da historiografia. Sem tentativa de romantizar, apenas sob a expectativa de observar a HD com uma autenticidade que lhe é

⁵⁵ Originalmente, “No one person can be expected to understand such a wide range of disciplines brought together in the digital humanities in the depth required to develop the innovative insights and methods that are the promise of the field.”

cabida, tanto quanto com a possibilidade de ser simbiótica à outras vertentes da pesquisa histórica. Dito isto, é perceptível que, com a chegada da digitalização do universo midiático, novas formas de fazer e redigir a história foram invariavelmente surgindo, como Lucchesi mostra ao escrever sobre hipertextualidade. Hipertextos hipermidiáticos são uma realidade constante e, a partir do momento que nos acostumamos com a ideia, confortáveis.

A HD traz consigo a possibilidade e a edificação de dispositivos computacionais enquanto auxiliares do trabalho de pesquisa e produção científica. Tais dispositivos podem ser *softwares*, linguagens de programação, instrumentos metodológicos (como o *crowdsourcing* ou os *bots* de inteligência artificial), algoritmos, etc. Nesse ínterim, ingressamos num campo muito mais amplo do que simplesmente a história que se espera digital: as humanidades digitais; o espaço de atuação onde diversos profissionais de inúmeras áreas se encontram; a grande “casa” repleta de cômodos que são as disciplinas científicas presentes neste universo multidisciplinar. As humanidades digitais, portanto, estão inseridas em um espaço de discussão da sua própria concepção: o que são? Um campo de estudo? Uma área de atuação? Uma comunidade de práticas? De que forma podemos citá-las é difícil estabelecer com certeza, mas sabe-se que o exercício consignado a elas é benéfico tanto às ciências humanas quanto às ciências computacionais.

As humanidades digitais agregam experiências conjuntas que coexistem para conseguir dar conta de demandas existentes a partir do advento do *big data* e da abundância de informações, elementos materiais e vestígios digitais. A partir disso, surgem projetos imersivos sob tecnologias que acompanham a historiografia; são projetos que utilizam do conhecimento de cientistas computacionais, informacionais e profissionais de outras áreas em consonância com o método científico das ciências humanas, que muito tem a contribuir para qualquer questão que envolva o ser humano, costumes e culturas (digitais ou não), bem como todo o processo da operação historiográfica que De Certeau cita.

Para isso, além de estarmos em colaboração com outras áreas, precisamos compreender que, enquanto expoentes de um campo **multidisciplinar**, não cabe a nós, se essa exigência/necessidade não existir, trabalharmos sozinhos. A multidisciplinaridade acontece justamente sob a perspectiva de estarmos em colóquio com outros profissionais. A atmosfera construída nesse ambiente de pluralidade acadêmica é espaço característico dessa comunidade.

Porém, enquanto trabalhamos sozinhos/as, ainda nos cabe o uso daquelas ferramentas digitais que nos servem como auxiliares tecnológicos importantes. Este trabalho fala justamente do uso, por parte de um/a profissional das ciências humanas, de ferramentas,

softwares e linguagens de programação que assistem seus trabalhos e contribuem enquanto facilitadores. Neste momento, ressalto a apreensão de que não houve nem a possibilidade, tampouco a ousadia, de mostrar e estudar todas as facetas que o tema escolhido apresenta. Ao explorar grafos, linguagens de programação e redes sociais virtuais, deixamos de lado muitos espaços de atuação possíveis - por óbvio, certos recortes são imprescindíveis. Logo, acho importante pontuar algumas coisas: que as linguagens de programação são inúmeras, cada uma servindo a um propósito e, dessa forma, não se estendem tão somente à raspagem e manipulação de dados; que, mesmo ao falar de Python, circunscrevi-me a um espaço muito específico do uso da linguagem - assim reitero que ela é ampla, suas bibliotecas são abundantes e, portanto, seu uso também é; que a raspagem de dados também foi escolhida de forma específica, sob o uso do *Twint* mas, certamente, com adequada procura, há outras formas de fazê-lo. Além disso, sobre o uso do Gephi, também me restringi a um uso mais basilar dele, mas, a partir dos exemplos utilizados, espero que tenha ficado inteligível a complexidade que ele pode atingir.

Tanto o Gephi quanto os dados são importantes atribuições a trabalhos multimodais, que pretendam alcançar o público leitor de outras formas. Produzir grafos, como mostrado no capítulo quatro, insere a pesquisa em um determinado espaço de didática que nos é favorável - para dizer o mínimo - e leva ao público uma nova forma de consumo do nosso material. São maneiras de visualizar a escrita da história que aproxima-se da história pública, ao mesmo tempo que do debate público. É mais uma forma de inserção dos/as profissionais da história nesses ambientes não-acadêmicos, mas de disputas narrativas e políticas muito importantes para serem simplesmente negligenciadas. Além disso, claro, servirão os grafos à historiografia como forma de restabelecer um diálogo com as próprias interpretações que temos somente tendo acesso aos dados brutos, produzindo diferentes acessos ao que eles têm a mostrar.

Ainda sobre a expressa função dos grafos à história, adiciono que a descrição feita neste trabalho sobre sua criação e utilização, inicialmente - carecendo do espaço e do tempo para uma análise específica e extensa - percebo-os como *quase*⁵⁶ um novo procedimento metodológico no campo da historiografia. Feitos a partir dessa interdisciplinaridade entre

⁵⁶ Entendo os grafos como dispositivos que, à história, servem como fonte primária ou, a depender, análise de fontes. A discussão deve ser sobre qual é o tratamento que o/a pesquisador/a dará aos dados e como interpretará esses dados. Se dados = fonte primária e grafos = tratamento destas fontes, então, os grafos são uma metodologia de análise de fonte; entretanto, para mim, dados e grafos apresentam *formas diferentes* de fontes, ambas capazes de serem analisadas. Sob essa ótica, ambos, portanto, são possíveis fontes primárias. Sobre serem um novo procedimento metodológico, acredito que esta assertividade requer um estudo um pouco mais amplo e mais complexo sobre a própria afirmação, realidade que não podemos alcançar nesta curta monografia. Mas, a priori, os resultados apresentados neste trabalho são promissores e contribuem para uma conclusão aprazível sobre grafos serem sim uma nova forma metodológica de perceber e pesquisar a historiografia.

programação e história, esses grafos são diferentes formas de narrativa, não-textual e aberta, bem como com informações não-listadas e, por vezes, subjetivas. É, para historiadores/as, uma linha interpretativa cuja existência será validada pelo método científico e visualizada pelo público conforme a argumentação do/a próprio/a historiador/a.

Mais importante, talvez, seja a concepção, através de tudo isso - as linguagens de programação, os grafos, os dispositivos digitais e todo o uso que fazemos deles -, das fontes nascidas digitais e qual sua aplicação à realidade dos/as pesquisadores/as de história. A mim é nítido que as FNDs são objetos de análise e estudo tais quais os vestígios históricos relativos à antiguidade clássica ou à idade média. Sua existência permite aos e às historiadores/as a produção de projetos e pesquisas que as utilizem enquanto fonte primária ou secundária, sendo possível compreendê-las enquanto o agente principal de seus trabalhos. As fontes nascidas digitais, portanto, ocorrem na e da internet, e fornecem à academia, à sala de aula e aos espaço de estudo uma amplitude no debate, na atuação e na concepção da própria teoria da história. Também ampliam o acesso aos saberes de diferentes culturas, de diferentes lugares, bem como a atuação de grupos sociais e grupos políticos, minorias ou não, no ciberespaço, possibilitando os estudos sobre a memória, os costumes, os comportamentos, etc. dessas comunidades. Factualmente, mesmo que, por vezes, “fonte primária” seja visualizada enquanto um manuscrito físico, antigo e carregado por uma autoridade de época, percebo-a, no estudo do tempo presente inserido no espaço digital, como aquela que fornece os acessórios necessários para fundamentar uma pesquisa que se inicia agora. Não obstante, indago se uma autoridade que produz, manipula e interpreta um conjunto de dados, um grafo ou quaisquer outras fontes nascidas digitais pode ser vista, ela própria, como a que carrega esses registros.

Por conta disso, apesar de formalmente discordar da ideia de que a HD serve mais à história pública do que à “pesquisa que faz argumentos”, ainda acredito que tanto a HD quanto às humanidades digitais e tudo o que foi citado neste trabalho referindo-se a elas, sejam catalisadores da história pública, principalmente no espaço em que entrecruzam os ‘muros da universidade’, ou simplesmente abrem caminhos para que o público de fora tenha acesso às produções que estão sendo feitas lá (ou aqui) dentro. Sobre essa percepção, reclino-me sob a *História Pública Digital* de Serge Noiret.

Dito isso, evidentemente é de minha expectativa que cada vez mais trabalhos e pesquisadores/as dediquem-se às possibilidades da história digital e tragam novas perspectivas ao campo da teoria da história, bem como aos correlatos campos de pesquisa de outras áreas da história. Tenho plena confiança que as linguagens de programação ainda

reservam muitas possibilidades, pesquisas, monografias de conclusão de curso, teses e dissertações. Anseio poder ter contato com trabalhos que explorem ainda mais a viabilidade do Python e do R, descobrindo novas funcionalidades e bibliotecas enquanto, eu mesmo, sigo nos estudos dessas linguagens. Acredito que também há um espaço em aberto muito importante para lidar com os usos das redes sociais virtuais na história; as narrativas presentes nelas são infundáveis, bem como as análises possíveis. O *Twitter* é um excelente espaço para o estudo das disputas de narrativa, principalmente pós-2016, com a eleição de Donald Trump e, 2018, com Jair Bolsonaro, dois presidentes que utilizaram integralmente das redes sociais como canal de comunicação oficial. Essas personalidades públicas fazem, constantemente, a história e, cabe a nós, estar presentes nesses debates. Mas, além do *Twitter*, há o *Instagram* (ainda, até onde eu tenho conhecimento, pouco explorado do ponto de vista histórico) e o *YouTube* (este tendo sido objeto de estudo do historiador Bruno Grigoletti Laitano), cujas ciberculturas criam uma atmosfera própria.

Penso, ainda, sobre as outras inúmeras linguagens de programação existentes no cosmo das ciências computacionais: o que elas nos oferecem? Podem nos auxiliar assim como o *Python*? Se sim, de que formas? Desejo poder estudar e observar outros/as pesquisadores/as se debruçando sobre pontos como esses. Igualmente, além do Gephi, quais são os outros *softwares* e/ou sites, ferramentas de código aberto, etc. que podem nos oferecer subsídios teórico-metodológicos? Por certo, temos diversas dessas já em uso por parte de humanistas: *Palladio*, *Tainacan*, *ArcGis*, *machine learnings* e inteligências artificiais, entre outros. Ainda assim, creio que podemos - e devemos - avançar nessas questões no Brasil, para nos posicionarmos como importantes agentes na esfera do debate dessas temáticas, no mundo, tanto quanto os Estados Unidos e a Itália, por exemplo, com seus campos referentes à HD bem estabelecidos (e constantemente servindo como pilares das nossas pesquisas). Almejo, dessa forma, perceber cada vez mais pesquisadores/as brasileiros/as estabelecidos/as nas humanidades digitais tanto quanto na história digital.

Em síntese, sob as limitações cabíveis a uma monografia de conclusão de curso, espero que a pretensão de contribuir para as discussões presentes ao longo do corpo textual deste trabalho tenha sido cumprida, de forma a acrescer à prateleira dessas ideias novas formas de pensar a teoria da história. Que com as reflexões feitas sobre os campos teóricos ou sobre as ferramentas seja possível colaborar para o progresso contínuo da pesquisa acadêmica e dos usos da internet. Encaro tais limitações, sem medo de soar quimérico, como um convite aos e às colegas para a continuidade desses estudos, mantendo um sensível olhar ao futuro e aos impactos que podemos fazer no mundo.

Finalmente, que se estenda neste trabalho a possibilidade de produção, cada vez maior, da escrita de uma história mais democrática, participativa e responsável com o mundo ao redor, seguindo o que Lucchesi, Silveira e Nicodemo argumentam ao dizer que “diante da escalada dos negacionismos que vivemos hoje, o texto de Castro é uma janela para respirarmos e pensarmos, juntos e em público, como nós historiadores podemos atuar na era digital para garantir o respeito à memória, a busca pela verdade e a justiça social” (LUCCHESI; SILVEIRA; NICODEMO, 2020, p. 164). Reitero as palavras dos próprios autores e digo que nunca fomos tão úteis.

6. REFERÊNCIAS BIBLIOGRÁFICAS

ALMEIDA, F. C. de. O Historiador e as Fontes Digitais: uma visão acerca da Internet como fonte primária para Pesquisas Históricas. **AEDOS**, [s. l.], v. 3, n. 8, p. 22, 2011.

ALVES, D. As Humanidades Digitais como uma comunidade de práticas dentro do formalismo acadêmico: dos exemplos internacionais ao caso português. **Ler História [Online]**, [s. l.], v. 69, 2017. Disponível em: <https://journals.openedition.org/lerhistoria/2496>.

AQUINO, I. **Análise de redes sociais**. [S. l.], [s. d.]. Portal. Disponível em: http://lhs.unb.br/cliomatica/index.php/An%C3%A1lise_de_redes_sociais. Acesso em: 3 set. 2021.

ARNOLD, T.; TILTON, L. New Data? The Role of Statistics in DH. In: DEBATES IN DIGITAL HUMANITIES (2019). MN: University of Minnesota Press, 2019. Part 3, Cap. 25. Online. Disponível em: <https://dhdebates.gc.cuny.edu/read/untitled-f2acf72c-a469-49d8-be35-67f9ac1e3a60/section/a2a6a192-f04a-4082-afaa-97c76a75b21c#ch24>

BASTIAN, M.; HEYMANN, S.; JACOMY, M. Gephi: An Open Source Software for Exploring and Manipulating Networks. *Em*: PROCEEDINGS OF THE THIRD INTERNATIONAL ICWSM CONFERENCE, 2009. **Anais [...]**. [S. l.: s. n.], 2009. Disponível em: <https://gephi.org/publications/gephi-bastian-feb09.pdf>.

BOLICK, C. M. DIGITAL ARCHIVES: DEMOCRATIZING THE DOING OF HISTORY. **International Journal of Social Education**, [s. l.], p. 122–134, 2006.

BORGMAN, C. L. Whose text, whose mining, and to whose benefit?. **Quantitative Science Studies**, [s. l.], v. 1, n. 3, p. 993–1000,

CARVALHO, B. L. P. de. História Pública e Redes Sociais na Internet: Elementos iniciais para um debate contemporâneo. **Revista Transversos**, [s. l.], v. 7, n. 7, p. 35–53, 2016.

CARVALHO, D. T. A. de. História ponto a ponto: o uso de ferramentas digitais para o acesso a fontes históricas. **VI ENECULT - encontro de estudos multidisciplinares em cultura**, [s. l.], p. 10, 2010.

CAUVIN, T. A. Ascensão da História Pública: Uma Perspectiva Internacional. **Revista NUPEM**, [s. l.], v. 11, n. 23, p. 8–28, 2019.

COHEN, D. J.; ROSENZWEIG, R. **Digital history: a guide to gathering, preserving, and presenting the past on the Web**. [S. l.]: University of Pennsylvania Press, [s. d.].

DE CERTEAU, Michel. **A operação historiográfica**. In: A escrita da história. Rio de Janeiro: Forense, 2013. p. 45-111;

FICKERS, A. Towards a New Digital Historicism? Doing History in the Age of Abundance. **VIEW Journal of European Television History and Culture**, [s. l.], v. 1, n. 1, p. 9, 2012.

FORTES, A.; ALVIM, L. G. M. Evidências, códigos e classificações: o ofício do historiador e o mundo digital. **Esboços: histórias em contextos globais**, [s. l.], v. 27, n. 45, p. 207–227, 2020.

FOSTER, M. Online and Plugged In?: Public History and Historians in the Digital Age. **Public History Review**, [s. l.], v. 21, p. 1–19, 2014.

GALLINI, S.; NOIRET, S. La historia digital en la era del Web 2.0. Introducción al dossier Historia digital. **História Crítica**, [s. l.], v. 43, p. 16–37, 2011.

GRANDJEAN, M. A social network analysis of Twitter: Mapping the digital humanities community. **Cogent Arts & Humanities**, [s. l.], p. 14, 2016.

JONES, S. E. The Emergence of the Digital Humanities (as the Network Is Everting). In: **DEBATES IN DIGITAL HUMANITIES (2016)**. MN: University of Minnesota Press, 2016.
Part 1, Cap. 1. Online. Disponível em:

<<https://dhdebates.gc.cuny.edu/read/untitled/section/09efe573-98e0-4a10-aaa3-e4b222d018fe#ch01>>

KIRSCHENBAUM, M. The .txtual condition: Digital Humanities, born-digital archives, and the future literary. **Digital Humanities Quarterly**, [s. l.], v. 7, n. 1, 2013. Disponível em: <http://www.digitalhumanities.org/dhq/vol/7/1/000151/000151.html#>.

_____. What Is Digital Humanities and What's It Doing in English Departments? In: DEBATES IN DIGITAL HUMANITIES (2012). MN: University of Minnesota Press, 2012. Part 1, Cap. 1. Online. Disponível em: <<https://dhdebates.gc.cuny.edu/read/untitled-88c11800-9446-469b-a3be-3fdb36bfb1e/section/f5640d43-b8eb-4d49-bc4b-eb31a16f3d06#en0102>>

KOKENSPARGER, B. **Guide to Programming for the Digital Humanities: Lessons for Introductory Python**. Omaha, NE: Springer, 2018. (SpringerBriefs in Computer Science).

LAITANO, B. G. **Digitalizar o arquivo, arquivar o digital : a história e suas fontes diante das velhas e novas tecnologias**. 2021. 108 f. Dissertação - Universidade Federal do Rio Grande do Sul, Porto Alegre, 2021. Disponível em: <https://lume.ufrgs.br/handle/10183/229967>.

LEME LOPES, A. P. Virada digital? Pesquisa histórica no ciberespaço. **Tempo e Argumento**, [s. l.], v. 10, n. 24, p. 136–169, 2018.

LIMA E FONSECA, T. N. de. Mídias e divulgação do conhecimento histórico. **Aedos**, [s. l.], v. 4, n. 11, p. 129–140, 2012.

LONDOÑO, F. Entrevista concedida a Isabel Jungk. **TECCOGS – Revista Digital de Tecnologias Cognitivas**, [s. l.], n. 21, p. 12–23, 2020.

LUCCHESI, A. Do texto ao hipertexto: notas sobre a escrita digital da história no século XXI. **VIII JORNADAS DE HISTORIA MODERNA Y CONTEMPORÁNEA. ENCUENTROS ENTRE LA POLÍTICA, LA ECONOMIA, LA CULTURA Y LA SOCIEDAD**, [s. l.], p. 11, 2012.

_____. Por um debate sobre História e Historiografia Digital. **Boletim Historiar**, [s. l.], v. 2, p. 45–57, 2014.

LUCCHESI, A.; SILVEIRA, P. T. da; NICODEMO, T. L. Nunca fomos tão úteis. **Esboços: histórias em contextos globais**, [s. l.], v. 27, n. 45, p. 161–169, 2020.

MALINI, F. Quando tudo parecia ser tão distante daqui: a eclosão das narrativas sobre covid-19. LABIC - Laboratório de Estudos sobre Imagem e Cibercultura. 2020. Disponível em: <<https://www.labic.net/cartografia/quando-tudo-parecia-ser-distante-daqui/>>. Acesso em: 19 de abr. 2020.

MOURA, M. A. Ciência da Informação e humanidades digitais: mediações, agência e compartilhamento de saberes. **Perspectivas em Ciência da Informação**, [s. l.], v. 24, n. especial, p. 57–69, 2019.

NOIRET, S. História Pública Digital | Digital Public History. **Liinc em Revista**, [s. l.], v. 11, n. 1, p. 28–51, 2015.

REGATTIERI, L. L. *et al.* A Forma Perspectiva no Twitter: uma técnica quanti-qualitativa para estudos de Redes Sociais. **XXXVII Congresso Brasileiro de Ciências da Comunicação**, [s. l.], p. 15, 2014.

ROBERTSON, S. The Differences between Digital Humanities and Digital History. In: DEBATES IN DIGITAL HUMANITIES (2016). MN: University of Minnesota Press, 2016. Part 4, Cap. 25. Online. Disponível em: <<https://dhdebates.gc.cuny.edu/read/untitled/section/ed4a1145-7044-42e9-a898-5ff8691b6628>>

ROLLAND, D. Internet e história do tempo presente: estratégias de memória e mitologias políticas. **Tempo**, [s. l.], v. 16, p. 59–92, 2004.

SANTAELLA, L. O metabolismo digital das humanidades. In: Humanidades digitais: performatividades na cultura digital [Ebook]. Goiânia: Cegraf UFG, 2021. p. 2-11. Online. Disponível em: <<https://publica.ciar.ufg.br/ebooks/invencoes/livros/9/capitulos/c01.html>>

SANTHIAGO, R. História pública e autorreflexividade: da prescrição ao processo. **Revista Tempo e Argumento**, [s. l.], v. 10, n. 23, p. 286–309, 2018.

SILVEIRA, P. T. da. **História, técnica e novas mídias: reflexões sobre a história na era digital**. 2018. 375 f. Doutorado - Universidade Federal do Rio Grande do Sul, Porto Alegre, 2018. Disponível em: <https://lume.ufrgs.br/handle/10183/189249>.

SINCLAIR, S.; RUECKER, S.; RADZIKOWSKA, M. **Information Visualization for Humanities Scholars**. [S. l.], 2013. Portal. Disponível em: <https://dlsanthology.mla.hcommons.org/information-visualization-for-humanities-scholars/>.

Acesso em: 3 set. 2021.

TELLES, H. V. HISTÓRIA DIGITAL, SOCIOLOGIA DIGITAL E HUMANIDADES DIGITAIS: Algumas questões metodológicas. **Revista Observatório**, [s. l.], v. 3, n. 5, p. 74–101, 2017.