

Tempo & Argumento

e-issn 2175-1803

O (re)nascer digital do arquivo histórico: investigações metodológicas do arquivo digital em *The Archived Web*

Resenha da obra:

BRÜGGER, Niels. *The Archived Web: Doing History in the Digital Age*. London: the MIT Press, 2018.

 Daniela Linkevicius de Andrade

Doutoranda em História na Universidade de Brasília (UnB).
Brasília, DF - BRASIL

lattes.cnpq.br/8334755280046247

dani.linkevicius@gmail.com

 orcid.org/0000-0002-4502-145X

Para citar esta resenha:

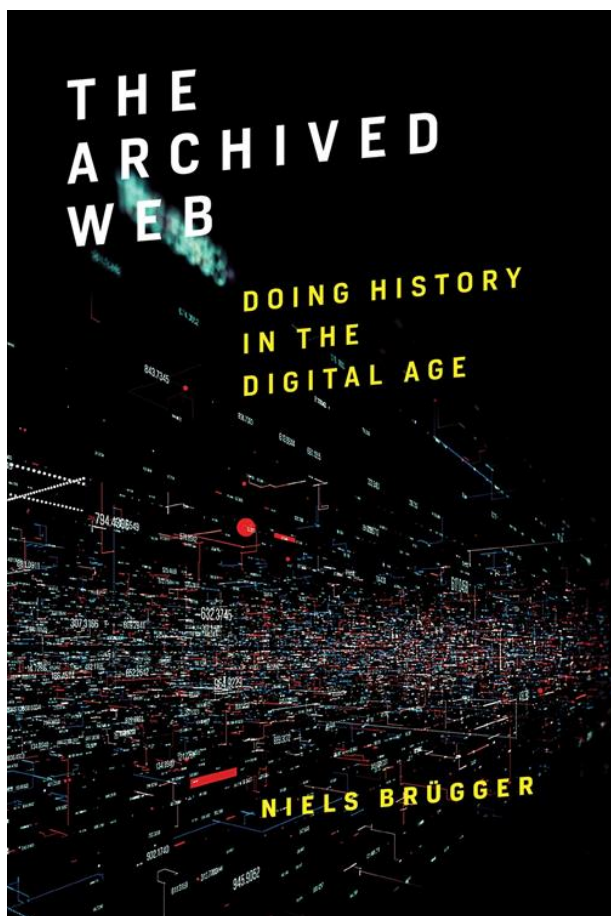
BRÜGGER, Niels. *The Archived Web: Doing History in the Digital Age*. London: the MIT Press, 2018. Resenha de: LINKEVICIUS, Daniela. O (re)nascer digital do arquivo histórico: investigações metodológicas do arquivo digital em *The Archived Web*. *Revista Tempo e Argumento*, Florianópolis, v. 12, n. 31, e0502. set./dez. 2020.

 <http://dx.doi.org/10.5965/2175180312312020e0502>

Recebido: 22/10/2020

Aprovado: 07/12/2020





Resenha da obra:

BRÜGGER, Niels. *The Archived Web: Doing History in the Digital Age*. London: the MIT Press, 2018.

Desde o final do século XX, cada vez mais pesquisadores se dedicam a compreender a formação de novas estruturas de sociabilidade online, que afetaram profundamente as relações sociais, culturais e políticas, com inúmeras implicações para a vida cotidiana. A partir dos anos 1990, a popularização do uso da *World Wide Web* nos colocou diante de um panorama em que uma das fontes

históricas mais importantes para o estudo da História do Tempo Presente, é, justamente, a web. Isso se dá não só através da compreensão do passado da própria web, como ferramenta para navegação na Internet, mas também da investigação do passado através da web, de forma que ela seja utilizada como fonte para outros objetos de estudo, como por exemplo conflitos globais, manifestações identitárias etc., que se entrelaçam com a web.

A pesquisa com fontes digitais, entretanto, ainda enfrenta muitos desafios que residem principalmente na pouca atenção metodológica que lhes foi dada. Estão, assim, intimamente relacionados a dois elementos principais: a especificidade de sua materialidade (ou seja, que é *digital*) e arquivamento (processo que envolve o momento em que a web está online até finalmente se transformar em uma coleção da web, pronta para ser investigada). Nesse sentido, a obra do dinamarquês Niels Brügger, *The Archived Web: Doing History in the Digital Age*, publicada em 2018 e ainda sem tradução para o português, tem como objetivo justamente oferecer uma estrutura teórica e metodológica que investiga

tais questões fundamentais relacionadas ao uso da web arquivada como uma fonte para os estudos históricos.

Brügger é um nome importante no que se refere a arquivos digitais. Professor do departamento de Estudos de Mídia na Aarhus University, sua produção é voltada a discussões dentro das áreas de Humanidades Digitais e História da Web, além do estabelecimento de arquivos digitais nacionais. Em *The Archived Web*, o autor faz sentir sua experiência ao longo de 10 capítulos.

O primeiro capítulo situa a História da Web, a qual autor defende ser aquela que pressupõe que a web do passado é digna de ser estudada como ponto de partida de uma pesquisa historiográfica. Ao situar a História da Web dentro do campo mais amplo da História Digital, destaca-se a importância de reconhecer que mais e mais fontes históricas vêm apenas em formato digital, o que deve levar os historiadores a reconsiderar suas teorias e métodos para alinhá-los a esse novo ambiente de origem.

Já o segundo capítulo inicia uma reflexão teórica através do que pode ser entendido por "digital", introduzindo, assim, o conceito de "digitalidade". Esses são fatores essenciais porque, conforme o autor, nem todas mídias digitais são digitais da mesma maneira (BRÜGGER, 2018, p. 21). Ao entender suas diferenças e similitudes, é possível agrupá-las e definir como serão analisadas. Nesse sentido, digitalidade refere-se às maneiras específicas pelas quais os bits digitais são materializados e combinados em textos e artefatos de mídia concretos.

Esse capítulo também apresenta as distinções entre material digital digitalizado, *born digital* e *reborn digital*, juntamente com a digitalidade específica da web online. Os materiais digitalizados são aqueles que existiram, anteriormente, em um formato não digital, mas que foram transformados para tornarem-se digitais. Por sua vez, o material *born-digital* é aquele que nunca existiu em nenhum outro formato que não o digital. Ele foi criado e disponibilizado apenas para mídia digital, como a rede de computadores por exemplo. Finalmente, o material denominado por Brügger como *reborn digital* é a fonte *born-digital* que foi coletada, preservada e, acima tudo, alterada no processo de arquivamento de maneira que não podemos mais considerá-la idêntica ao material *born-digital* do qual é proveniente, até porque, em muitos

casos, esse “original” é efêmero e pode ter sido editado, atualizado e/ou excluído (BRÜGGER, 2018, p. 22).

O capítulo três completa a base teórica para as reflexões metodológicas sobre o uso da web arquivada como fonte, iniciada nos capítulos anteriores. Ele sistematiza uma grade analítica de pesquisa da web, dividida em cinco extratos. Esses extratos são: o elemento web (qualquer entidade semiótica na forma de elemento escrito, som, imagem estática ou em movimento); a página web (uma única janela do navegador); o site (unidade analítica composta por páginas da web inter-relacionadas); esfera da web (material da web relacionado a um tópico, tema, evento ou área geográfica); e a web como um todo — fenômenos que transcendem outros estratos, como os algoritmos de busca, o navegador, etc. (BRÜGGER, 2018, p. 35).

Após essa exposição, o quarto capítulo fornece uma variedade de casos ilustrativos dos últimos 15 anos de estudos de História da Web. Cada caso é um exemplo de como a escrita da História da Web foi conduzida e como as questões de historiografia foram abordadas por estudos de História da Web existentes.

No capítulo cinco, Brügger assume o ponto de vista do indivíduo ou organização que assume a tarefa de preservar a web. O capítulo começa perguntando por que é necessário arquivar a web online, considera se um arquivo da web é, de fato, um arquivo e apresenta uma definição ampla de arquivamento da web. O arquivamento da web deve incluir coleta e preservação, o que significa um processo deliberado e proposital — isto é, o ato de arquivar deve ser acompanhado por algum tipo de consciência de que se está realmente arquivando a web (BRÜGGER, 2018, p. 79-80).

O autor ainda identifica cinco formas de arquivamento na web: (1) criar uma imagem, (2) fazer um filme em tela, (3) baixar arquivos individuais, (4) *web crawling*, (5) coletar material da web de um banco de dados, disponibilizado através de uma interface de programação de aplicativos — API, (6) coletando a web que foi retirada off-line e preservada inalterada e (7) coletando a web conforme apresentado em outros tipos de mídia, como livros, filmes e televisão.

O sexto capítulo, por sua vez, assume a perspectiva do pesquisador que deseja estudar a web arquivada, preservada através de um acervo já existente e disponibilizado por instituições especializadas, como é o caso do Internet Archive, e por arquivos nacionais, como o Arquivo.pt (arquivo da web portuguesa) e Danish Netarkivet (arquivo da web dinamarquesa). Portanto, o capítulo oferece um panorama dos principais tipos de coleções, onde procurar a web do passado e o que se pode esperar encontrar em cada caso.

O capítulo sete aborda a próxima etapa no processo de pesquisa, examinando mais de perto como a web do passado se apresenta como um objeto de estudo nos diferentes tipos de coleções da web, incluindo o mapeamento dos principais desafios colocados pelas diferentes digitalidades da web arquivada. O autor destaca algumas características fundamentais desse tipo de arquivo que necessitam ser levadas em consideração. Dentre elas, podemos citar a ausência de um documento original que possibilite consultas para conferir a qualidade do que foi arquivado, o que ocasiona a incompletude das fontes *reborn digital*. Isso dificulta a determinação do que antes estava faltando e por quê (BRÜGGER, 2018, p. 105). Por esse motivo, a fonte *reborn digital* é uma versão única e não uma cópia do que antes era online. A mesma entidade online arquivada no passado por dois pesquisadores diferentes pode vir a ser versões diferentes, em vez de cópias idênticas do que estava online no momento do arquivamento (BRÜGGER, 2018, p. 107).

O oitavo capítulo traz considerações acerca de como um acadêmico pode usar a web arquivada em projetos de pesquisa, reintroduzindo os cinco estratos da web com o objetivo de discutir os desafios e possibilidades que o pesquisador encontra ao buscar estudar a web em cada um dos estratos pelo uso dos diferentes tipos de coleções.

O capítulo nove apresenta como a digitalidade da web arquivada impacta algumas das questões práticas relacionadas à pesquisa de História da Web. Para o autor, o conceito de digitalidade implica na necessidade de reinterpretação e tradução das abordagens metodológicas tradicionais, principalmente se considerarmos a necessidade de avaliação das várias versões de um mesmo

documento e maneiras de referenciar a fonte para se adequar às novas condições.

Por fim, o último capítulo debate as interseções entre a História da Web e a história de outras mídias digitais e apresenta alguns dos desafios que a História da Web enfrentará futuramente devido à falta de fontes digitais e à pouca preocupação com o arquivamento desse tipo de documento. Nos encaminhamos, assim, para a conclusão da obra, em que o autor descreve alguns itens possíveis de serem incluídos em uma futura agenda de pesquisa de História da Web.

A conceitualização e sistematização realizada por Brügger pode parecer simples, mas não é. A escassez de discussão metodológica sobre fontes digitais frequentemente ignora a materialidade específica de fontes *born digital* e/ou sequer mencionam sua transformação em acervo documental, tornando-se *reborn digital*. Mesmo discussões dentro da História Digital podem se tornar frustrantes àqueles que desejam trabalhar com os diferentes extratos analíticos da web. Isso porque a web não costuma aparecer como fonte ou objeto de estudo, mas como plataforma de trabalho para pesquisadores que comumente trabalham com fontes digitalizadas. Mesmo historiadores que trabalham com História Digital frequentemente não têm habilidade para lidar criticamente com a web arquivada.

A frustração pode levar a uma dúvida primordial no trabalho de pesquisa: por onde começar a analisar tais fontes digitais, quando minha intenção não é (somente) fazer uso de softwares ou desenvolver novas ferramentas digitais para a disciplina histórica, mas entender o passado com a web e através da web? Nessa perspectiva a obra de Brügger pode aparecer como uma opção viável para começar a entender essa fonte e evitar erros metodológicos básicos, como por exemplo analisar um site em sua versão online, e não arquivada, que podem influenciar enormemente nos resultados encontrados da investigação. Além disso, ao sugerir diversos tipos de estratégias de arquivamento, *The Archived Web* é capaz de incluir desde pesquisadores mais versados em linguagens de programação, bem como aqueles com pouco nenhum domínio.

Em um momento em que a História do Tempo Presente se esforça continuamente em adaptar-se às novas linguagens digitais, os historiadores ainda parecem receosos do universo da World Wide Web. Mas, entender a realidade do seu estudo, demandará não somente que dominem o caos das bibliotecas e acervos históricos tradicionais, como que se joguem cada vez mais naquilo que Brügger chama de World *Wild* Web.