Sobre Arquivos, Big Data e Pesquisa Histórica

About Files, Big Data and Historical Research

Resumo

A intenção deste artigo é colocar a questão da importância dos arquivos para o trabalho do pesquisador. Há inúmeras maneiras de traduzir o processo de conhecimento, e uma delas é estabelecer relações de desconfiança e estranhamento com aquilo que enxergamos a olho nu, e é a partir do trabalho com os documentos e a produção de provas que o estatuto científico da história vem sendo discutido nos últimos tempos. E isso se dá à medida que o empirismo, espécie de contrapeso à retórica pós-modernista, ganha força entre os seus praticantes.

Palavras-Chave: Arquivos, Pesquisa Histórica, Big-Data, Documentos, Hiperlinks

Abstract

The intention of this article is to pose the question of the importance of archives to the researcher's work. There are innumerable ways of translating the process of knowledge, and one of them is to establish relationships of distrust and estrangement with what we see with the naked eye, and it is from the work with the documents and the production of evidence that the scientific statute of history has been discussed in recent times. And this happens as empiricism, a kind of counterbalance to postmodernist rhetoric, gains strength among its practitioners.

Keywords: Files, Historical Search, Big-Data, Documents, Hyperlinks

Prof.^a Dr.^a. Ana Maria de Almeida Camargo.

Professora sênior do
Departamento de História da
Faculdade de Filosofia, Letras
e Ciências Humanas da
Universidade de São Paulo.

Interpretar é sair da superfície das coisas, é vê-las à distância, é criar descontinuidades, é contornar o que está dito e posto, é encarar o outro... Há inúmeras maneiras de traduzir o processo de conhecimento, e uma delas é estabelecer relações de desconfiança e estranhamento com aquilo que enxergamos a olho nu. A despeito do que se possa considerar mais importante nesse processo - e tudo tem importância, na verdade -, é a partir do trabalho com os documentos e a produção de provas que o estatuto científico da história vem sendo discutido nos últimos tempos. E isso se dá à medida que o empirismo, espécie de contrapeso à retórica pós-modernista, ganha força entre os seus praticantes¹.

O assunto não é novo, e muito já se escreveu a seu respeito. Os clássicos manuais de introdução aos estudos históricos sempre concederam grande espaço à heurística e à postura crítica diante dos documentos. Um dos mais prestigiados atualmente, o de Antoine Prost², reitera que a referência a fontes primárias e secundárias é constitutiva da própria metodologia empregada pelos historiadores. Se a contiguidade entre pesquisa e documentação acabou por se transformar em conversa para principiantes, tal o efeito da simbiose que se reconhece entre os termos desse binômio, ganha agora fôlego redobrado com a crescente expansão do mundo digital.

Não basta parafrasear o documento, reproduzi-lo no todo ou em parte, ou indicar as séries das quais se extraíram dados transformados em gráficos e percentuais de frequência. A figura do historiador como mediador entre as fontes utilizadas e o leitor de sua obra não se sustenta mais nesses antigos pilares, sempre colocados, aliás, sob suspeita: na ótica do senso comum, pelo menos, a prova documental fica invariavelmente associada a leituras enviesadas e a exclusões de conveniência. Refutá-las (à maneira de Karl Popper) implicaria refazer o mesmo caminho do pesquisador, visitar os

mesmos arquivos e consultar os mesmos documentos... Como isso é possível?

Criou-se agora uma nova “cultura” documental. As notas de rodapé vêm sendo gradativamente substituídas por *hiperlinks*, que permitem acesso direto ao texto integral das obras consultadas, e já se pensa na adoção obrigatória de plataformas digitais com as fontes primárias e os dados efetivamente obtidos na pesquisa, para efeitos de compartilhamento³. Paralelamente à massiva digitalização de todos os tipos de documentos, vive-se hoje a experiência das Humanidades Digitais, que tem servido de parâmetro para diferentes iniciativas nessa área. Tudo leva a crer que estamos no melhor dos mundos, a desfrutar do pleno acesso que a tecnologia nos oferece de modo ininterrupto e surpreendente.

É preciso considerar, no entanto, não apenas as peculiaridades da pesquisa histórica, tão decantadas quanto negligenciadas, mas as que envolvem o conceito de arquivo, pouco lembradas e igualmente postas de lado.

Quando Daniel Fabre⁴ afirma que o historiador mantém uma relação especial com as fontes, ao contrário do que se observa em outras disciplinas, chama a atenção para uma dessas peculiaridades. Perguntar a um psicólogo, a um sociólogo ou mesmo a um geógrafo quais são suas fontes, pondera o autor, faz com que discorram sobre os procedimentos de pesquisa adotados e sobre o modo como elaboram as perguntas que vão propor à realidade observada. Antes desse questionamento, as fontes

1 Veja-se, a respeito, DAVIES, Stephen. *Empiricism and history*. London: Palgrave Macmillan, 2003.

2 PROST, Antoine. *Objectivité, vérité, preuve*. In: - **Douze leçons sur l'histoire**. Paris: Seuil, 1996. p. 288-293.

3 Na esteira das principais agências estrangeiras de fomento, a Fapesp (Fundação de Amparo à Pesquisa do Estado de São Paulo) passou a exigir, desde 2017, ao menos no âmbito dos projetos temáticos, o compromisso dos pesquisadores com a divulgação dos dados levantados (utilizados ou não). A perspectiva é propiciar o desenvolvimento científico e a racionalização dos investimentos feitos em diferentes áreas de pesquisa. Ver PIERRO, Bruno de. Uma estratégia para dados: pesquisadores são estimulados a gerenciar e compartilhar as informações científicas que produzem. **Pesquisa Fapesp**, São Paulo, n. 267, maio 2018.

4 FABRE, Daniel. *L'ethnologue et ses sources*. *Terrain - Anthropologie et Sciences Humaines*, Paris, n. 30, p. 3-12, mar. 1988.

tos passa a assumir, então, um outro e fundamental papel: o de fornecer enquadramento e limites (em outras palavras, contexto e balizamento) para o uso de seu conteúdo manifesto.

Figura 2

Fonte: www.escolakids.uol.com.br

O historiador, por sua vez, vale-se em grande medida de documentos que, acumulados nos arquivos, não decorrem das hipóteses que levanta, e sim das atividades exercidas por pessoas físicas e jurídicas (públicas e privadas) no desempenho de suas funções. Obedientes a lógicas distintas daquelas que os historiadores imprimem a suas pesquisas, tais documentos oferecem a seus usuários secundários uma vantagem única: o fato de constituírem provas das ações que lhes deram origem. E é a partir desse estatuto probatório dos documentos de arquivo, que incide única e exclusivamente sobre o contexto imediato em que foram produzidos, que o historiador poderá, ele também, instituir os documentos que lhe permitirão praticar seu ofício. Se admitimos que o historiador tem o dom de converter qualquer coisa em documento, como se depreende da leitura de Michel de Certeau ⁵, pode e deve exercê-lo também sobre os documentos de arquivo -- os documentos de “nascença”, como diria Marie-Anne Chabin⁶. O caráter probante originário de tais documen-

Fonte: <http://www.simi.org.br/>

No âmbito das Humanidades Digitais e dos *Big Data*, em que qualquer conjunto documental recebe o nome genérico de arquivo sem que disponha de organicidade ou outros atributos distintivos, prevalece o tom otimista com que Jo Guldy e David Armitage⁷ caracterizaram o trabalho do historiador ante as possibilidades que ora se anunciam. Como não aderir ao incomensurável volume de dados disponíveis, à existência de ferramentas capazes de condensá-los e torná-los visíveis (como Zotero, Paper Machines, ManyEyes), às possibilidades de abordagem transnacional e ao alargamento de horizontes temporais (a longa duração de Braudel colocada, enfim, ao alcance de todos)? Como não aderir à perspectiva de abandonar os minúsculos e paroquiais recortes que dominam hoje nossa produção historiográfica?

O poder de sedução de projetos de espectro tão variado -- voltados para grandes temas, como a Guerra de Secessão (*The Valley of the Shadow*), para séries relativamente homogêneas, como as que cobrem o tráfico transatlântico de escravos do século XVI ao XIX (*The Trans-Atlantic Slave Trade Databases*), para personagens que se destacaram por suas qualidades, individualmente ou em grupo

5 “Em história, tudo começa com o gesto de separar e reunir, transformando-os em documentos, determinados objetos antes dispostos de outra maneira”. CERTEAU, Michel de. *L'écriture de l'histoire*. Paris: Gallimard. 1975. p. 84. (Bibliothèque des Histoires).

6 Para Marie-Anne Chabin, os “documentos de nascença” não se confundem com os “de batismo”, isto é, aqueles que recebem estatuto probatório por força das operações a que são submetidos por juristas, historiadores e outros interessados, no propósito de fundamentar sentenças, teses e decisões. Ver, a respeito: CHABIN, Marie-Anne. Tout est archive. In: - Je pense donc j'archive: l'archive dans la société de l'information. Paris: L'Harmattan, 1999. p. 37-69.

7 **The History manifesto.** Cambridge: Cambridge University Press, 2014.

(*The William Blake Archive* e *Women Writers Project*) e tantos outros -- não exime as Humanidades Digitais das críticas que lhes são feitas⁸. A principal delas é, como no caso da chamada digitalização “selvagem”, a ausência de metadados relacionados com a proveniência dos documentos, procedimento que não pode ser entendido, de forma simplista e apressada, como menção a organismos de custódia.

Trata-se aqui do contexto de produção, das marcas que caracterizam a funcionalidade original dos documentos acumulados ao longo das atividades (inclusive as de pesquisa) de entidades e pessoas diversas. De tais marcas dependerão, certamente, duas coisas importantíssimas: um efetivo e fecundo compartilhamento de dados e a viabilidade do exercício crítico a que se propõem, afinal, os próprios historiadores.

REFERENCIAS

BERRY, David M.; FAGERJORD, Anders. **Digital humanities: knowledge and critique in a digital age**. Cambridge: Polity Press, 2017.

CERTEAU, Michel de. **L'écriture de l'histoire**. Paris: Gallimard, 1975.

CHABIN, Marie-Anne. *Tout est archive*. In: - **Je pense donc j'archive: l'archive dans la société de l'information**. Paris: L'Harmattan, 1999. p. 37-69.

DAVIES, Stephen. **Empiricism and history**. London: Palgrave Macmillan, 2003.

FABRE, Daniel. *L'ethnologue et ses sources*. **Terrain - Anthropologie et Sciences Humaines**, Paris, n. 30, p. 3-12, mar. 1988.

PIERRO, Bruno de. Uma estratégia para dados: pesquisadores são estimulados a gerenciar e compartilhar as informações científicas que produzem. **Pesquisa Fapesp**, São Paulo, n. 267, maio 2018.

PROST, Antoine. *Objectivité, vérité, preuve*. In: - **Douze leçons sur l'histoire**. Paris: Seuil, 1996. p. 288-293.

The History manifesto. Cambridge: Cambridge University Press, 2014.

TRELEANI, Matteo. **Mémoires audiovisuelles: les archives en ligne ont-elles un sens?** Montréal: Les Presses de l'Université de Montréal, 2014. (Parcours Numériques).

8 A exemplo das que se leem em: BERRY, David M.; FAGERJORD, Anders. **Digital humanities: knowledge and critique in a digital age**. Cambridge: Polity Press, 2017. E também em: TRELEANI, Matteo. **Mémoires audiovisuelles: les archives en ligne ont-elles un sens?** Montréal: Les Presses de l'Université de Montréal, 2014. (Parcours Numériques).