

Exploring legacies of race and slavery in our historical information environment: text analysis of the Encyclopaedia Britannica

Charlton, Ash

a.r.charlton@sms.ed.ac.uk

University of Edinburgh, United Kingdom

Introduction

This research uses text analysis to interrogate early editions of *Encyclopaedia Britannica* (*EB*) in exploring explicit and implicit references to slavery through editions to understand how this changes diachronically. Combining close and distance reading methods, it raises important methodological questions about how we analyze legacies in historical information environments.

The first eight editions of the *EB* were published from 1768-1860, covering a period from the height of the transatlantic slave trade to abolition of the slave trade within the British Empire. As a popular reference publication that gathered knowledge of science and the arts which has an enduring legacy itself, these early editions offer an insight into how *EB* (and the society it served) portrayed and reinforced concepts of race and slavery. Using text analysis, this research interrogates our understanding of race and slavery in *EB* by exploring explicit and implicit links made through each edition, and how these change (or do not) diachronically. In doing so, it raises issues of how digital methods can be used to understand legacies of enslavement, and possibilities for decoloniality, within our information environment.

History

Although individual entries in *EB* are often cited in historical research, the extent and depth of these editions mean that little attempt has been made to analyze them at scale. The first eight editions of *EB* comprise of 109,393 pages, from 135 volumes. Jeff Loveland's 2019 study of European encyclopaedias claimed that "computers' ability to recognize and interpret patterns of text remains limited" (Loveland 2019: 13). Despite this claim, in recent years there have been several encyclopaedia-focused digital humanities projects, including research on citation strategies in *Encyclopédie* (Edelstein et al. 2019), topic modelling of *Encyclopédie* to benefit research in classifications systems (Roe et al. 2016) and a study of automatic nineteenth century *EB* entry indexing (Grabus et al. 2019). These advances demonstrate that digital approaches offer possibilities for novel interpretation of information sources at scale.

The many references to race and slavery within *EB*'s pages are little-explored and although some entries have been quoted by

researchers (Sebastiani 2013: 169, and Popkin, 1974: 257-258), they are otherwise difficult to connect. Vincent Brown has argued that addressing our historical records of slavery requires "rigorous and responsible creativity" and that archives are not simply the records themselves but encompass our research methods and tools and how researchers choose to visualize and share results (Brown 2015: 134). This research seeks to rigorously make connections through *EB* via creative digital means and make the results available in an accessible and visually engaging way.

Method of inquiry

Open-access and reusable data from the National Library of Scotland's *Data Foundry* provides the basis of this research, using images from the *EB* dataset with handwritten text recognition platform Transkribus (READ-COOP 2022) to create a high-quality machine-readable dataset with fewer character errors than the original OCR (National Library of Scotland 2019). The "openness, transparency and practicality" of these datasets offer a firm opportunity for collaboration between the library and digital researchers (Ames & Lewis 2020: 3). The resulting texts are being analyzed via a combination of close reading of encyclopaedia entries with quantitative digital analysis methods. This research primarily uses R within RStudio to conduct the text analysis and generate engaging data visualizations. Encyclopaedia entry headings extracted using the humanities toolbox for text-based archives, *defoe*, will also be used in conjunction with the analysis to generate network maps of connections in the text (Filgueira et al. 2019; 2021).

This research uses topic modelling to closely examine the entries relating to slavery and enslavement to offer an overview of the concepts and how they change. The identification and frequency of selected keywords across all entries in editions offers an insight into intellectual interconnectivity between *EB* entries, both within singular editions and broader comparisons across multiple editions. These keywords are broadly grouped into three categories: commodities and trade, geographic terms, and identifying terms for people or organizations. By identifying which entries keywords appear in, and whether these are consistent across multiple editions, the connections between encyclopaedia entries will be mapped to form an intellectual network. In doing this, the absence of keywords in entries where they may be expected, and the absence of entries, will also be examined to reveal legacies of race and slavery in *EB*.

This research contribution has significant implications for several disciplines, especially the history of slavery, and encyclopaedia studies. It raises important methodological questions as to how we begin to develop best practice approaches for recognizing and analyzing the legacies within our information environment.

Bibliography

Ames, Sarah / Lewis, Stuart (2020): "Disrupting the Library: Digital Scholarship and Big Data at the National Library of Scotland", in: *Big Data & Society* 7, 2: 1-7 <https://doi.org/10.1177/2053951720970576> [14.12.2021].

Brown, Vincent (2015): "Mapping a Slave Revolt: Visualizing Spatial History through the Archives of Slavery", in *Social Text* 33, 4: 134-41 <https://doi.org/10.1215/01642472-3315826> [27.06.2022].

Edelstein, Dan / Morrissey, Robert / Roe, Glenn (2013) "To Quote or Not to Quote: Citation Strategies in the *Encyclopédie*",

in *Journal of the History of Ideas* 74, 2: 213–36 <https://www.jstor.org/stable/43291299> [15.08.2022].

Filgueira, Rosa / Jackson, Michael / Roubickova, Anna / Krause, Amrey / Ahnert, Ruth / Hauswedell, Tessa / Nyhan, Juliane / Beaven, David / Hobson, Timothy / Ardanuy, Mariana Coll / Colavizza, Giovanni / Hetehrington, James / Terras, Melissa (2019): "Defoe: A Spark-Based Toolbox for Analysing Digital Historical Textual Data", in *2019 15th International Conference on EScience (EScience)*: 235–42 <https://doi.org/10.1109/eScience.2019.00033> [21.09.2022].

Filgueira, Rosa / Grover, Claire / Karaiskos, Vasileios / Alex, Beatrice / Van Eyndhoven, Sarah / Gotthard, Lisa / Terras, Melissa (2021) "Extending Defoe for the Efficient Analysis of Historical Texts at Scale", in *2021 IEEE 17th International Conference on EScience (EScience)*: 21–29 <https://doi.org/10.1109/eScience51609.2021.00012> [21.09.2022].

Grabus, Sam / Greenberg, Jane / Logan, Peter / and Boone, Jane (2019): "Representing Aboutness: Automatically Indexing 19th- Century Encyclopedia Britannica Entries", in *NASKO* 23: 138–48 <https://doi.org/10.7152/nasko.v7i1.15635> [03.08.2022].

Loveland, Jeff (2019): *The European Encyclopaedia from 1650 to the Twenty-First Century*. Cambridge: Cambridge University Press.

National Library of Scotland (2019): "Encyclopaedia Britannica", in *Data Foundry* <https://data.nls.uk/data/digitised-collections/encyclopaedia-britannica/> [28.10.2021].

Popkin, Richard (1974): "The Philosophical Basis of Eighteenth-Century Racism". in *Studies in Eighteenth-Century Culture* 3: 245–62. <https://doi.org/10.1353/sec.1974.0016> [27.09.2021].

Roe, Glenn / Gladstone, Clovis / Morrissey, Robert (2016): "Discourses and Disciplines in the Enlightenment: Topic Modeling the French Encyclopédie", in *Frontiers in Digital Humanities* 2: 1-13 <https://www.frontiersin.org/articles/10.3389/fdigh.2015.00008> [03.08.2022].

Sebastiani, Silvia (2013): *The Scottish Enlightenment: Race, Gender, and the Limits of Progress*. New York: Palgrave Macmillan.

READ-COOP (2022): *Transkribus*: <https://readcoop.eu/transkribus/> [16.01.2022].