Legacies of race and slavery in our historical information environment: exploring knowledge networks in the *Encyclopaedia Britannica*

Ash Charlton
@AshRCharlton

25th July 2023
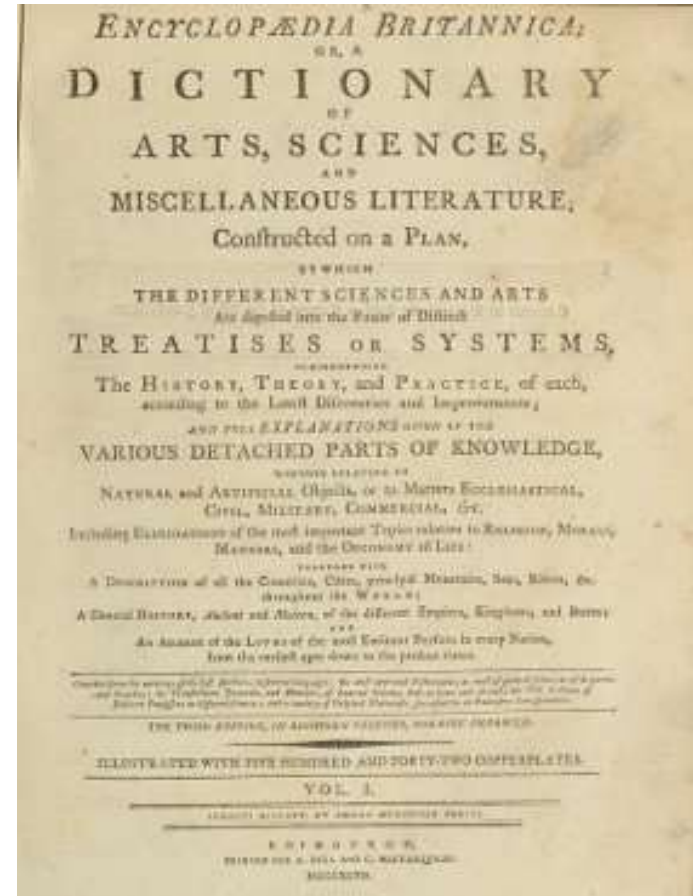ISKO UK Conference

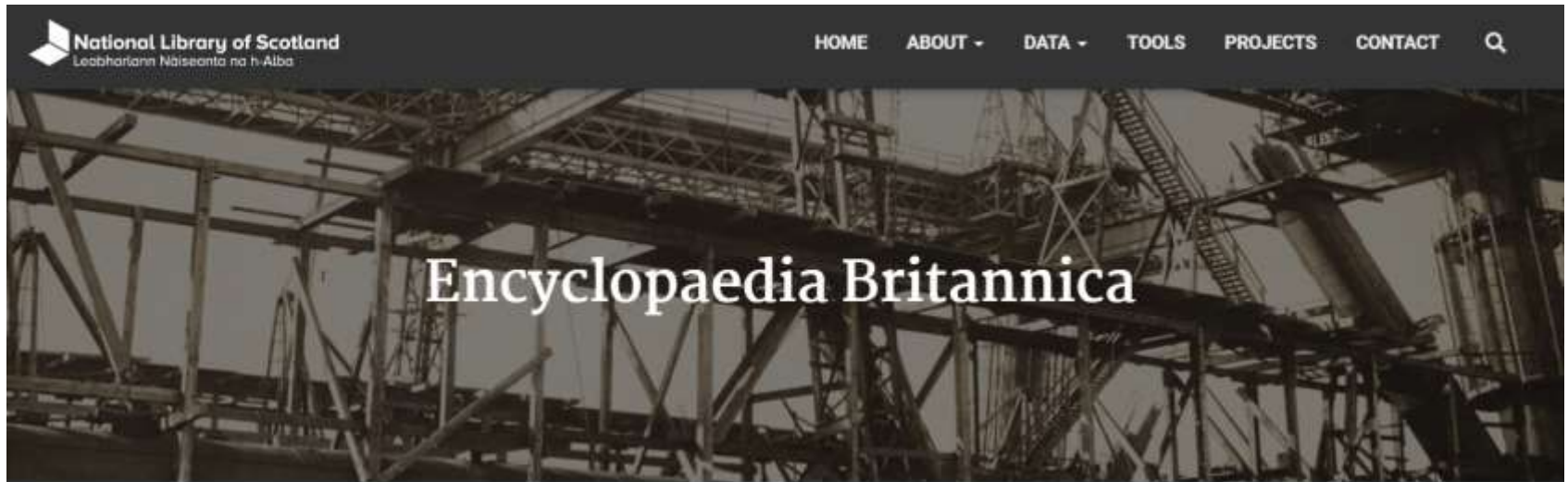Influencing the world since 1583

Key questions

- How can patterns of change in subject knowledge be identified over time?

- What types of knowledge networks are formed between terms and articles in *Encyclopaedia Britannica*?

Encyclopaedia Britannica

- First eight editions

- 1768-1860

- 109,393 pages

- 135 volumes

THE UNIVERSITY *of* EDINBURGH

National Library of Scotland
Leabharlann Nàiseanta na h-Alba

HOME   ABOUT ▾   DATA ▾   TOOLS   PROJECTS   CONTACT   Q

Encyclopaedia Britannica

Original OCR: no clean-up

155,388 ALTO XML files at page level

155,388 image files

METS metadata files at item level

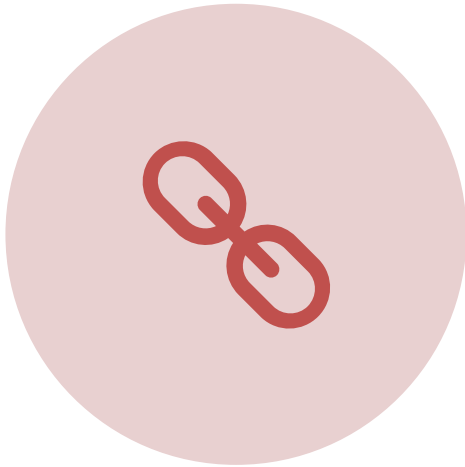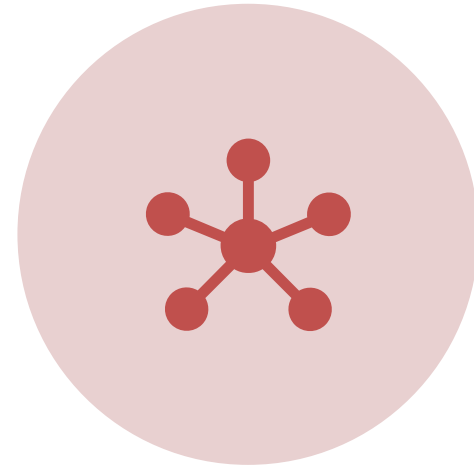19,257,785 lines and 166,729,009 words

Covers years 1768-1860

https://data.nls.uk/data/digitised-collections/encyclopaedia-britannica/

Influencing the world since 1583

Datasets

- National Library of Scotland's Data Foundry
- https://data.nls.uk/data/digitised-collections/encyclopaedia-britannica/

- Nineteenth Century Knowledge Project
- https://tu-plogan.github.io/

- Transkribus (using NLS data)

Information types

EXPLICIT

IMPLICIT

Articles with expected information sources

SLAVE, a perfon in the abfolute power of a mafter, either by war or conqueft. We find no mention of flaves before the deluge: but immediately after, *viz.* in the curfe of Canaan: whence it is eafily inferred, that fervitude increafed foon after that time; for in Abraham's time we find it generally eftablifhed.

COTTON, in commerce, a foft downy fubftance found on the bombax, or cotton tree. See BOMBAX.

Cotton is feparated from the feeds of the plant by a mill, and then fpun and prepared for all forts of fine works, as flockings, waiftcoats, quilts, tapeftry, curtains, &c. With it they likewife make muflin, and fometimes it is mixed with wool, fometimes with filk, and even with gold itfelf.

The fineft fort comes from Bengal and the coaft of Coromandel.

JAMAICA, an ifland of America, fituated in the Atlantic ocean, between 76° and 79° of weft longitude, and between 17° and 18° odd minutes north latitude, near 5000 miles fouth-weft of England, 100 miles fouth of the ifland of Cuba, and 350 miles north of Terra Firma. The ifland lies eaft and weft, and is about 140 miles long, and 60 broad. The wind fets on the fhore almoft all the day in every part of the ifland, and off the fhore in the night; it fometimes hails, but the people there never fee froft or fnow.
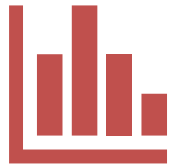
SUGAR, in natural biftory, is properly the effential falt of the fugar-cane, as tartar is of the grape. See CHEMISTRY p. 161. and SACCHARUM.

This plant rifes to eight, nine, or more feet high; the ftalk,

'Slave', 'Cotton', 'Jamaica' and 'Sugar' articles from the first edition

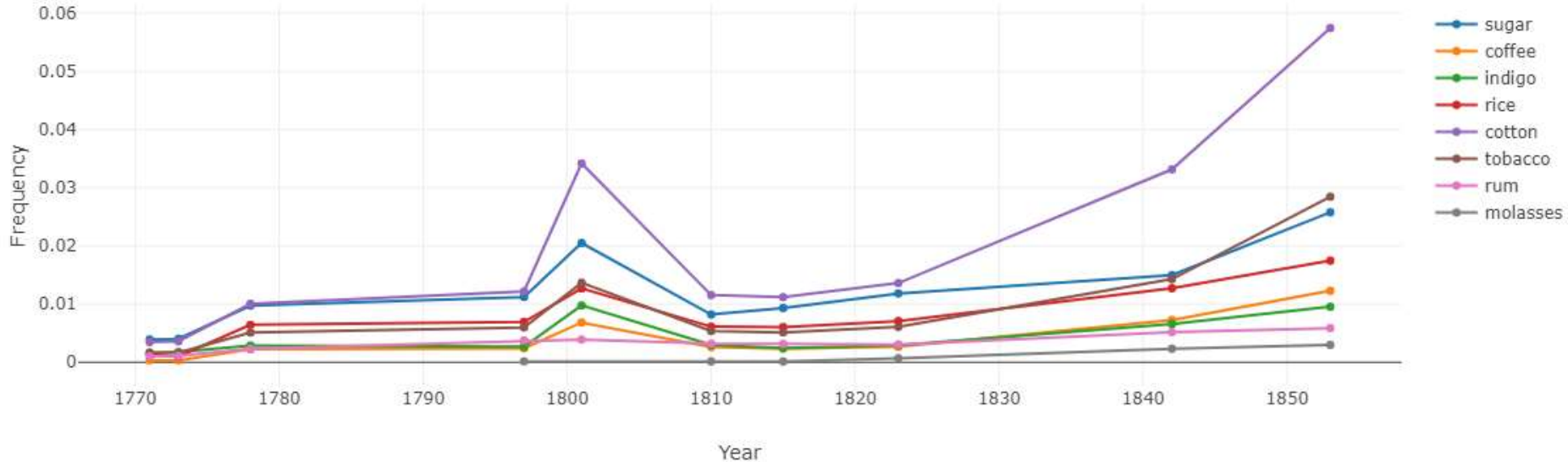# Text mining

Frequency counts   Keyword context   Network analysis

Frequency counts with *frances*

Normalized Frequency of Lexicon Term per Years

Forthcoming article: *frances: Cloud-based Historical Text Mining with Deep Learning and Parallel Processing* (2003) Rosa Filgueira, Lilin Yu, Wilfrid Askins, Melissa Terras & Ash Charlton
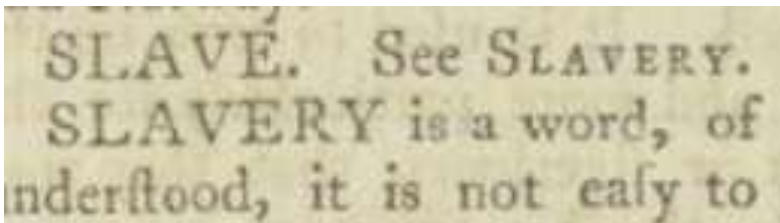
Keywords in context

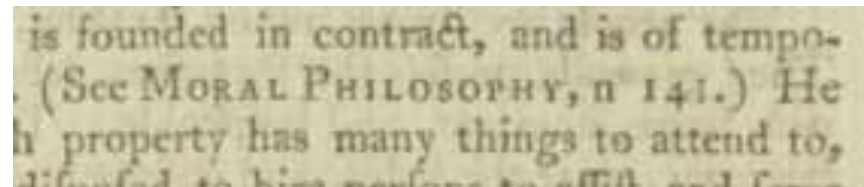| Edition	Term	Entry	Context
First, Vol 1	slaves	Addicti	Ancient Rome
First, Vol 1	slaves	Addicti	Ancient Rome
First, Vol 1	slaves	Alligati	Ancient Rome
First, Vol 1	slaves	Antheresia	Ancient Rome
First, Vol 1	slaves	Bagnio	Contemporary
First, Vol 1	slaves	Bastonade	Ancient Rome
First, Vol 1	slaves	Bestiarii	Ancient Rome
First, Vol 1	slaves	Burning	Ancient Rome

Network analysis

- Which articles are linked explicitly with reference points? e.g. 'See X'

- Network maps to visualise these connections

- Implicit connections?

Explicit article links in the third edition

'Slave' article links in 3rd edition

Slave

Slavery

Philology | Right | Moral Philosophy | Gladiators | Property | Theology

Conclusion & next steps

- Text mining as a valuable tool in information collection

- Extracted data can be used to identify patterns in knowledge and to map knowledge organisation in *EB*

- Wider applications in constructed information sources

References

- Filgueira, R., Yu, L., Askins, W., Terras, M., Charlton, A. *frances: Cloud-based Historical Text Mining with Deep Learning and Parallel Processing* (2003) [forthcoming, IEEE Science Conference]

- National Library of Scotland. 'Encyclopaedia Britannica – Data Foundry', 2019. https://data.nls.uk/data/digitised-collections/encyclopaedia-britannica/

- 'Nineteenth-Century Knowledge Project'. Accessed 4 August 2022. https://tu-plogan.github.io/

- READ-COOP. 'Transkribus'. Accessed 8 February 2022. https://readcoop.eu/transkribus/

Thank you!

@AshRCharlton

A.R.Charlton@sms.ed.ac.uk