Trading Consequences

Project Team: York University (Prof Colin Coates, Dr. Jim Clifford), the University of Edinburgh (Prof Ewan Klein, Dr Beatrice Alex, Dr Claire Grover), University of St. Andrews (Prof Aaron Quigley, Dr. Uta Hinrichs) and EDINA (James Reid and Nicola Osborne)

Overview of goals and objectives of Trading Consequences

This project has combined two leading-edge areas of historical enquiry (transnational history and environmental history) with innovative techniques developed in geo-referencing and visualisation. The team of historians, computational linguists and computer scientists has explored ways of analysing and visualising global trade in commodities in the 19th century and the associated environmental consequences. Using a variety of digitised texts, this project extracts information on primary resources harvested, mined, traded and consumed in the 19th century, geo-locates the place name mentions for these resources and correlates them with a set of potential environmental impacts. The general goal has been to adapt geo-referencing and visualisation techniques for historians researching in large quantities of textual data. For the computational linguists and computer scientists, historians have served as a test group for honing their techniques when applied to big data.

Challenges and lessons learned from international collaboration across disciplines and domains

The different tasks of the project have been divided across teams located in different countries (data-mining and visualisation in Scotland; historical interpretation and complementary research in Canada). Regular monthly Skype calls, as well as a few face-to-face meetings, mostly organised around workshops or conferences, have ensured on-going collaboration. One of the historians in the project has unusually advanced computer skills, and this has facilitated collaboration with other colleagues.

Access to data sets has occasionally been difficult to secure. In some cases, working in different countries has allowed members of the group to leverage access more successfully than if this had been a one-country project.

Digital humanities, social sciences and computational based research methods in the context of big data projects

We have achieved encouraging success in our initial research goals, but we have realised that we have had to modify our questions as we confront the limitations of the text collections available to us. The project has analysed and compared problems with existing digitized historical text collections: the OCR quality has varied a great deal between sources. Some of these issues can be overcome by automatic OCR post correction, for instance to correct false 'f' characters to 's' characters. Others are products of bad print and paper quality or early OCR projects. We overcome these OCR challenges by using ten million pages of digitised historical documents, allowing us to explore large trends in the history of commodities in the British world across the 19th century.

Indicators of success

We have presented interim results of this project to a variety of potential users and have benefitted from a range of feedback. In particular, we organised a workshop with a group of environmental historians. They have responded very positively to the heuristic potential of the public database that will emerge out of this study. We are in the process of comparing the text-mining results against more traditional historical techniques but have yet to collate the findings of this validation process.

<div align="center">**Trading Consequences**</div>

Measuring impact
The feedback from test audiences has been encouraging to date. We have established valuable links with a variety of individual researchers and groups interested in the history of commodity trades as a result of allowing users to test the tool or presenting our work to scholars in the humanities and social sciences.

Knowledge dissemination mechanism and tools
One of the main goals of this project is to create a database and web-based visual interface for researchers to explore these data and to advance their own research endeavours. We have disseminated Trading Consequences in papers or presentations at conferences (e.g. OR 2012, North American Conference on British Studies 2012, European Society for Environmental History 2013, Canadian Historical Association 2013), workshops (e.g. LTHist 2012, the AHRC Commodity Histories Project Networking Workshop in 2012, DemoFest 2012, the Digital Scholarship: day of ideas event in 2013) and two summer schools (CHESS 2013 and Big Data InfoVis 2013) and in a series of blog posts and guest blog post (for a complete list see the Trading Consequences Publications and Presentations page).
The Trading Consequences blog was set up at the start of the project in (Jan 2012) using WordPress with hosting supported by EDINA. It has been used to reach out both to the academic historical research community, to those interested in data mining and visualisation, and to audiences more broadly interested in the 19th century and the history of international trade. All blog posts have been written by members of the project team and have been crafted to be engaging, entertaining and accessible although each post carries real detail about the process of the evolving project or the historical research process. The blog has attracted over 2000 unique visitors making over 3300 visits to the site to date (Jan 2012-Aug 2013). Whilst most blog visitors have accessed the site from desktop machines (2967 visits to date) around 9% of visitors have accessed it via tablets and mobile devices (376 visits to date) reflecting the usefulness of ensuring a mobile-friendly version is available.
Twitter (@digtrade) has been used throughout the project to amplify various aspects of Trading Consequences. All blog posts have been promoted via tweets, several receiving flurries of Retweets and interest. Events, conferences and publications have also been amplified via Twitter, sometimes through tweets indicating attendance, presentations, etc. and sometimes through live tweeting, retweeting and reporting from those events. We have also received a number of compliments and responses to presentations via Twitter.

Importance of working with libraries, archives and data repositories
Our library contacts at York University have been particularly helpful in helping us liaise with data providers. The archives at Kew Gardens have also been very generous in providing data and an opportunity to post preliminary findings.

Capacity building and training (students and highly qualified personnel)
At York University, we have employed on the project three postdoctoral researchers, two PhD students, one MA students and three undergraduate students. One postdoctoral researcher obtained a tenure-track position in a History Department partway through the project, and he has continued to play a leading role in advancing the research. We anticipate that he will submit further research grant proposals to pursue this and related projects. Another researcher obtained a two-year postdoctoral fellowship.   At the University of Edinburgh, we have employed on the project one research fellow and one PhD student. The research fellow has taken on a central role in managing the project. This project has allowed her and the rest of the team to make new contacts with scholars in social sciences and digital humanities which are leading to new potential collaborations.   At St. Andrew's University, we have employed one postdoctoral researcher on a three-year contract extending beyond the duration of this project and two undergraduate students. The postdoctoral researcher has taken the lead on visualizing the mined data. She has organised a number of workshops and was instrumental in organising a Big Data InfoVis summer school attended by 35 PhD students from around the world. This summer school was concerned with the processing, management and presentation of "big data" in an intelligible form with information visualisation techniques and methods.