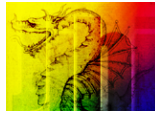


# JOURNAL OF INTERDISCIPLINARY HISTORY OF IDEAS



2019

Volume 8 Issue 15

Item 4

– Section 3: Notes –

Distant Reading, ‘The Great Unread’, and  
19<sup>th</sup>-Century British Conceptualizations of  
the Civilizing Mission

A Case Study

by

Darren Reid



## JIHI 2019

Volume 8 Issue 15

### Section 1: Editorials

1. *Editorial* (JIHI)

### Section 2: Articles

2. *Vouloir pour la nature. La représentation juridique des entités naturelles* (P. Brunet)
3. *On the lability of natural entities, at the example of Raspe's De novis e mari natis insulis* (E. Pasini)

### Section 3: Notes

4. *Dislant Reading, 'The Great Unread', and 19<sup>th</sup>-Century British Conceptualizations of the Civilizing Mission: A Case Study* (D. Reid)

### Section 4: Reviews

5. *Retro-Review: Michel Verdon, Keynes and the 'Classics': a study in language, epistemology and mistaken identities, 1996. Essay Review* (R. Bobulescu, M. Laudet)
  6. *Literature on Inheritance: A Summary of What Can Be Learnt* (C.G. Marian)
- .....

# Distant Reading, 'The Great Unread', and 19<sup>th</sup>-Century British Conceptualizations of the Civilizing Mission

## A Case Study

Darren Reid \*

*Franco Moretti's work on 'the great unread' raises serious questions for any field. In the field of new imperial history, I am especially interested in how the use of distant reading to access 'the great unread' can inform our study of imperial and colonial discourses. To stimulate research in this direction, I conducted a small-scale case study to explore the methodological considerations of applying distant reading to imperial discourse analysis. I performed a distant reading on a small sample of the British Library's collection of digitized nineteenth century monographs to trace how understandings of the idea of a 'civilizing mission' changed over the century. My findings indicate that a core understanding of the civilizing mission as related to agriculture, education, and peace remained constant throughout the century, while peripheral understandings of the civilizing mission appear to have changed from tangible social values such as commerce and rule of law to intangible values such as honesty and liberty. Based on the methodological choices I made, primarily due to the small scale of this case study, these findings should be approached with caution. However, they illustrate how probing 'the great unread' can generate new research questions and provide nuanced contexts of conceptual change in which to locate the study of civilizing discourses.*



\* University of Victoria ([dreid820@uvic.ca](mailto:dreid820@uvic.ca)).



In his “Conjectures on World Literature”, Franco Moretti borrowed Margaret Cohen’s term ‘the great unread’ to describe the huge mass of primary source material that is not analyzed by scholars. Moretti argued that the traditional method of close reading forced scholars to focus on a few texts, the canon, and prevented scholars from considering the vast amount of information that existed, the non-canon. He suggested that any scholarships produced upon such a limited source base could not be considered truly representative of reality. As a solution to this problem, Moretti defined a new method of “distant reading”, which replaces a zoomed-in study of singular sources with a zoomed-out study of many.<sup>1</sup> Moretti’s work on ‘the great unread’ raises, to my mind, serious questions for any field. In my field of new imperial history, I am especially interested in how the use of distant reading to access ‘the great unread’ can inform our study of imperial and colonial discourses. To stimulate research in this direction, as well as to serve as a preliminary test case in preparation for a future research project, I performed a small-scale case study to explore the methodological considerations of applying distant reading to imperial discourse analysis. As the subject of this case study, I performed a distant reading on the British Library’s collection of digitized 19<sup>th</sup>-century monographs to trace how understandings of the idea of a ‘civilizing mission’ changed over the century. I identified this as a subject worth distant reading because there is a debate in the historiography over what Victorians understood their “civilizing mission” to entail, with both sides of the debate limiting their source base in different ways. One school of thought uses canonical texts to argue that Victorians continued to believe in their civilizing mission through to the end of the 19<sup>th</sup> century. Another school uses a temporally restricted collection of canonical and non-canonical texts to argue that British civilizing discourses in the late 19<sup>th</sup> century consisted only of shallow platitudes designed to distract Britons from the realities of imperialism.

<sup>1</sup> Franco Moretti, “Conjectures on World Literature”, *New Left Review* 1 (2000): 54-68, <https://tinyurl.com/moretti2000conjwl>.

By distant reading civilizing discourses in the British Library's digitized collection, and thereby sidestepping the limited source bases of these two schools of thought, I hoped to gain new insights that might move this debate forward. This short paper presents the details of my case study. I begin by summarizing the historiographical debate over the meaning of the 19<sup>th</sup>-century British civilizing mission, and then presenting the distant reading methodology and how it can benefit the historiography of British civilizing missions. Next, I describe how I adapted the distant reading method to my case study and the methodological choices I made. I explore four methodological decisions that were fundamental to performing this distant reading, and which any distant reading project must take into account: choosing an exploratory or a model-based framework, choosing a corpus, choosing a periodization to organize data points, and choosing how to derive relevant meaning from short snippets of text. Finally, I give a brief interpretation of my findings, which indicate that a core understanding of the civilizing mission as related to agriculture, education, and peace remained constant throughout the century while peripheral understandings of the civilizing mission appear to have changed from tangible social values such as commerce and rule of law to intangible values such as honesty and liberty. Based on the methodological choices I made, primarily due to the small scale of this case study, these findings should be approached with caution. However, they illustrate how probing 'the great unread' can generate new research questions and provide nuanced contexts of conceptual change in which to locate the study of civilizing discourses.

## Historiography

In attempting to identify 19<sup>th</sup>-century British conceptualizations of their civilizing mission, historians have established two distinct approaches, which I term the comparative and the discursive. The comparative approach attempts to define British understandings of their civilizing mission through a comparative analysis of contemporary theoretical writings with first-hand accounts in order to reveal shared underlying understandings between the theorists and the practitioners. For example, consider the work of Jorgen Osterhammel and

John Comaroff. Jürgen Osterhammel argues that the civilizing mission was understood as a legal phenomenon that lasted from 1830-1870,<sup>1</sup> the purpose of which was to “make nations peaceful, warrior classes obsolete, and individuals industrious and acquisitive”.<sup>2</sup> To support this claim, he leans upon two types of primary sources. First, he uses theoretical texts by Nicholas de Condorcet (French philosopher, 1743-1794), Turgot (French economist, 1727-1781), and Adam Smith (Scottish philosopher, 1723-1790) to prove the existence of this conception of the civilizing mission. He then uses official documents such as legal statutes and parliamentary debates to prove the praxis of this conception. John Comaroff develops a framework of the civilizing mission similar to Osterhammel’s, in which he differentiates between the founding ideas and the practical implementation of the civilizing mission. Comaroff locates the understandings of the civilizing mission at the intersection of various theoretical and popular authors, specifically: “the philosophical support of Bentham and Mill [...] the artistic work of Shelley and Blake [...] and the political economy of Smith and Ricardo”.<sup>3</sup> However, while Osterhammel holds that the civilizing process was primarily implemented through legal work, Comaroff asserts that it was primarily implemented through missionary work, and uses the writings of missionaries such as David Livingstone and Robert Moffat as his sources.<sup>4</sup> From such missionary writings, Comaroff argues that the civilizing mission was conceived as a religious phenomenon that lasted from 1830-1900 and that had five purposes: to banish superstition, to transform the chaotic landscape into organized cities, to recast the gendered division of labour, to spread modern methods of agriculture, and to teach self-sufficiency through reading and writing.<sup>5</sup>

These two arguments use the same strategy to overcome the limited scope of close reading. Each uses similar types of primary sources: famous philosophers, important politicians, and highly-circulated missionary writers. These types of

<sup>1</sup> Jürgen Osterhammel, *Europe, the ‘West’ and the Civilizing Mission* (London: The German Historical Institute London, 2006), 20.

<sup>2</sup> *Ibid.*, 21.

<sup>3</sup> John Comaroff, “Images of Empire, Contest of Conscience: Models of Colonial Domination in South Africa”, in *Tensions of Empire: Colonial Cultures in a Bourgeois World*, ed. by Frederick Cooper and Ann Laura Stoler (Berkeley: University of California Press, 1997), 169.

<sup>4</sup> *Ibid.*, 181.

<sup>5</sup> *Ibid.*, 181-2.

sources can be considered the historiographical canon. No one source can represent a historical society, yet the close reading method of the comparative approach limits the capacity of a historian as to how many sources they can realistically read. To overcome this issue, the comparative historians have chosen to focus on the canon because it consists of sources that we know were highly influential and circulated widely within their contemporary society. While these canonical sources might have influenced upper-class conceptions of the British civilizing mission, they cannot be seen to represent the civilizing conception of the vast majority of British society. By relying on the canon, the comparative approach can only represent the past to the extent that these canonical sources represent their contemporary society.

The discursive approach to the history of British civilizing missions is, in part, a reaction to the comparative approach's focus on canonical texts. The discursive approach attempts to access 19<sup>th</sup>-century British understandings of their civilizing mission by identifying patterns of representation in civilizing discourses. For example, consider the work of Catherine Hall and Zoe Laidlaw. Zoe Laidlaw argues that, while civilizing language generally remained constant from 1837 onwards, a "series of silences" developed in the official texts of Britain and its colonies, where officials tried to draw attention away from what was actually happening to Indigenous peoples by celebrating past humanitarian victories such as the criminalization of slavery.<sup>1</sup> For Laidlaw, the Victorian civilizing mission existed in theory but not in practice, and was used as a discourse to distract attention "from the erasure—literal and metaphorical—of colonized peoples."<sup>2</sup> Her research draws upon a significantly larger corpus than the comparative histories. Instead of focusing on canonical texts, Laidlaw set out to study the records of all who "oversaw, endorsed and enabled its [the civilizing mission's] processes: the politicians and the civil servants; the land promoters and emigration agents; the supportive families of migrants; and the landed classes who paid the passages of their too-numerous tenants."<sup>3</sup> However, her research is still limited to the number of sources she could reasonably read. Her

<sup>1</sup> Zoë Laidlaw, "Imperial Complicity: Indigenous Dispossession in British History and History Writing", in *Emancipation and the Remaking of the British Imperial World*, ed. by Catherine Hall, Nicholas Draper, and Keith McClelland (Manchester: Manchester University Press, 2015), 131-2.

<sup>2</sup> *Ibid.*, 132.

<sup>3</sup> *Ibid.*, 143.

breadth of sources was greater than that of the comparative historians, but she was sharply limited in terms of time frame, and she recognizes that her work “confines itself to a short, if crucial, period in the mid-century”.<sup>1</sup> Catherine Hall takes a similar position as Zoe Laidlaw. She argues that civilizing discourses translated abstract notions of savagery-versus-civility into articulated “grammars of difference”.<sup>2</sup> She maintains that the Victorian civilizing mission was an idea that was constructed by British historians and popular writers as the 19<sup>th</sup> century progressed as a framework for understanding and justifying British imperial activity.<sup>3</sup> Much like Laidlaw’s claim that civilizing discourse was utilized to erase imperial oppression, Hall argues that civilizing discourse was utilized to reimagine imperial oppression. However, she also shares Laidlaw’s limited temporal scope. Her study focuses solely on primary sources published between 1857 and 1860, although it must again be recognized that she pulls from a much wider variety of primary source than the comparative histories.

In comparing these two approaches, we can see that each employs a different strategy for gaining representational authority, yet that each approach is limited by the number of sources that can be close read. The comparative histories depend on the representational authority of canonical texts, but by so doing they can only represent a narrow section of society. Conversely, the discursive histories draw representational authority from observing rhetorical patterns from a broader source base, but in broadening their source base they narrow their temporal scope. I do not mean to suggest that these limitations of the close reading method invalidate these approaches in any way. Neither the comparative nor the discursive histories claim to represent the entirety of British society, and when read together they complement each other well, one covering a greater temporal scope and one a greater source scope. However, by applying a distant reading method to this subject and thereby integrating ‘the great unread’ into the analysis, it is conceivable that new insights can be gleaned that would not otherwise be apparent through close reading. After all, the purpose of distant reading is to integrate non-canonical texts into schol-

<sup>1</sup> *Ibid.*, 132.

<sup>2</sup> Catherine Hall, *Civilizing Subjects: Colony and Metropole in the English Imagination, 1830-1867* (Chicago: University of Chicago Press, 2002), 53-60.

<sup>3</sup> *Ibid.*, 54.



arship that previously considered only the canonical. In his seminal work on distant reading, Moretti wrote:

The trouble with close reading (in all of its incarnations, from the new criticism to deconstruction) is that it necessarily depends on an extremely small canon. This may have become an unconscious and invisible premise by now, but it is an iron one nonetheless: you invest so much in individual texts only if you think that very few of them really matter.<sup>1</sup>

The historiography of civilizing missions may not be necessarily flawed, but the limit that the close reading method has placed upon the source suggests that there is information out there in ‘the great unread’ that might improve our knowledge of British conceptions of civilizing missions. Distant reading can help us access this knowledge, leading me to devise this case study. My goal was to create a timeline of how civilizing discourse changed throughout the entire 19<sup>th</sup> century. By doing so, I hoped to contextualize the claims of both the comparative and the discursive histories of 19<sup>th</sup>-century British understandings of the civilizing mission. For instance, the discursive claim that civilizing discourse after the 1850s turned from the intention to civilize to the erasure of oppression cannot be sustained without evidence that the way civilizing was talked about changed substantively from pre-1850 to post-1850. On the other hand, the comparative claim that civilizing missions were predominantly ‘about’ a particular set of values, whether those be legal, religious, or otherwise, cannot be sustained without evidence that those particular values were discussed more often than others. Traditional methods of close reading are poorly suited to gathering these types of evidence due to the limited number of sources than can feasibly be analysed, but distant reading is not so limited.



<sup>1</sup> Moretti, “Conjectures on World Literature”, 57.

## Designing the case study

In devising my case study, the first methodological consideration to be made was whether to employ a model-based framework or an exploratory framework. The original distant reading method, as described by Franco Moretti, was intended to use computational quantitative tools to create explanatory models of historical change over time. In his essay, “Style, Inc.: Reflections on 7,000 Titles (British Novels, 1740-1850)”, Moretti combined bibliographic lists of British novel titles to create a master bibliography of 7,000 titles from 1740-1850. He plotted the length of titles and the publication density of novels over time. From these graphs, Moretti suggested a simple model of how British novel titles changed over time: “The market expands, and titles contract”.<sup>1</sup> Another example of a model-based framework is Jean-Baptiste Michel et al., “Quantitative Analysis of Culture Using Millions of Digitized Books”. Their team plotted numerous data points, from the number of words per book to the frequency of individual words, and used these graphs to suggest a range of explanatory models. For instance, they interpret shortening time periods between the peak and the nadir frequencies of specific years, i.e. “1973”, to mean that as time progresses we forget history at an increasingly faster pace.<sup>2</sup> A fundamental aspect of these quantitative distant readings is an iterative process: explanatory models are derived from observable trends, and those models are continuously tested against further evidence until the model most accurately reflects reality.

A model allows the researcher to isolate aspects of an interesting phenomenon, and in discovering certain properties of such aspects, the researcher can continue revising the model to identify additional properties. In this “iterative process”, the “truth” of that phenomenon resides at some asymptomatic point that can never be reached. But along the way, the modeling process yields productive insights.<sup>3</sup>

This iterative process, “endlessly proceeding from the social reality to the model, and then back again”, in the words of Fernand Braudel, is crucial to

<sup>1</sup> Franco Moretti, “Style, Inc. Reflections on Seven Thousand Titles (British Novels, 1740-1850)”, *Critical Inquiry* 36, no. 1 (2009): 141.

<sup>2</sup> Jean-Baptiste Michel et al., “Quantitative Analysis of Culture Using Millions of Digitized Books”, *Science* 331, no. 6014 (2011): 179.

<sup>3</sup> Richard Jean So, “All Models Are Wrong”, *PMLA* 132, no. 3 (2017): 669.

the model-based framework, and for this reason I determined it was not an appropriate framework for this case study.<sup>1</sup> This case study was to be a small-scale exploration of the methodological considerations of distant reading, and in the limited timeframe I operated within, I would not have been able to follow through on the iterative process. I therefore chose to employ an exploratory framework of distant reading.

As surveyed by Stefan Jänicke et al., scholars have developed various exploratory distant reading frameworks, which use patterns within data points to identify new areas of research.<sup>2</sup> For instance, Mae Capozzi and Scott Enderle's "A Distant Reading of Empire" used topic modeling on books published in Britain between 1757-1795, and identified a pattern: references to India tended to increase around significant events in the American colonies, such as the Townshend Duties in 1767 and the Boston Tea Party in 1773. Based on their distant reading, they suggest that Britons considered the American colonies and India to be more related than was previously believed, and therefore that this is an important avenue for new research.<sup>3</sup> As opposed to model-based frameworks, which use distant reading to produce new historical arguments through iterative processes, exploratory frameworks don't aim to produce fully-fledged historical arguments, but rather use distant reading to suggest new areas of research. By so doing, exploratory frameworks can produce a valuable scholarly deliverable in a shorter amount of time, with the trade-off being that any deliverable produced from an exploratory distant reading will necessarily require more work. I deemed this framework most applicable for this case study, and project managers must also take into account the purpose and the length of their project in order to decide which framework is most appropriate for them.

For this case study, I designed an exploratory distant reading project that would plot the values applied by 19<sup>th</sup>-century British authors to their civilizing mission. The overall structure of my method followed three stages. First, I would identify individual manuscripts that included invocations of civilizing

<sup>1</sup> Fernand Braudel, "History and the Social Sciences", in *On History*, trans. by Sarah Matthews (Chicago: University of Chicago Press, 1980), 45.

<sup>2</sup> Stefan Jänicke et al., "On Close and Distant Reading in Digital Humanities: A Survey and Future Challenges", *EuroVis* (2015): 10-11, <https://doi.org/10.2312/eurovisstar.20151113>.

<sup>3</sup> Mae Capozzi and Scott Enderle, "A Distant Reading of Empire", Accessed August 28, 2019, <https://readingfromadistance.wordpress.com>.

discourses from within a larger corpus of British manuscripts, and divide these into three time periods by publication date: 1800-09, 1850-59, and 1890-99. Second, I would identify paragraphs which invoked civilizing discourses, and digitally annotate them for the abstract values they attributed to their civilizing mission. Third, I would plot these values over time and observe any changes that occurred.



After defining an exploratory framework, my second methodological consideration was choosing which corpus to analyze. It is best to start with identifying your specific needs, and then determining whether a corpus already exists that can meet those needs or whether you will have to build your own. For this case study, I identified two needs. First, I needed full-text reading capability. Many digitized manuscript corpuses, such as the HathiTrust Research Centre's Extracted Features datasets, are designed specifically for statistical linguistic analysis. Rather than containing full sentences, they merely contain metadata and word frequency data. For this case study, which seeks to identify changes in how civilizing missions were conceptualized by Britons throughout the 19<sup>th</sup> century, these corpuses are of little use. Word frequency data can show changes in how often the words "civilized" or "civilizing" was used over time, but without reading the sentences in which the words were used, we cannot know how these words were used or the values that were applied to these words. In order to identify these meanings, I needed a corpus that contained full-text data. Second, I needed accurate metadata. This is rarely an issue, since most digital collections are curated by library and research institutions which follow strict metadata standards, but it is a need worth noting. In any distant reading, and especially an exploratory distant reading, it is important to ensure that any extracted data points can be easily traced back to the original source. Otherwise, any researcher who wishes to pursue a newly generated research question will have no indication of where to start. Based on these needs, I chose the British Library's collection of 60,937 digitized 19<sup>th</sup>-century manuscripts, available at

<https://data.bl.uk/digbks/>. This collection is composed of manuscripts that were digitized in a joint British Library-Microsoft initiative started in 2006. The manuscripts chosen for digitization are described with the following characteristics:

- Predominantly written in English.
- Predominantly published in Britain.
- Predominantly published between 1800 and 1936.
- Relevant to the study of geography, history, English literature, and philosophy.
- Based on originals which are between 18.5 and 30.5 centimeters in height and not thicker than 10 centimeters.<sup>1</sup>

The manuscripts are available in ALTO XML format, which is a filetype that contains Optical Character Recognition data. ALTO XML files are highly useful for research, as they contain both full-text data as well as page image data. However, there are few freely available tools to view and manipulate ALTO XML files, meaning that some knowledge of XML parsing is generally required to use these files. This is another important consideration when planning a distant reading project. For this case study, I wrote my own Python scripts, which can be found at <https://github.com/dreid820/distantreading>, to convert these ALTO XML files into full-text XML files. All this means is that I converted text strings containing the location data of individual words on a page into full sentences without location data. For example, consider the ALTO XML phrase:

```
<SP ID="P43_SP00004" HPOS="233" VPOS="175" WIDTH="49" />
<String ID="P43_ST00006" HPOS="282" VPOS="144" WIDTH="33" HEIGHT="32"
  CONTENT="It" WC="0.76" CC="40" />
<SP ID="P43_SP00005" HPOS="315" VPOS="176" WIDTH="16" />
<String ID="P43_ST00007" HPOS="331" VPOS="145" WIDTH="24" HEIGHT="31"
  CONTENT="is" WC="0.83" CC="30" />
<SP ID="P43_SP00006" HPOS="355" VPOS="176" WIDTH="13" />
<String ID="P43_ST00008" HPOS="368" VPOS="145" WIDTH="116" HEIGHT="32"
  CONTENT="indeed" WC="0.79" CC="660000" />
```

<sup>1</sup> Adrian Edwards, “19<sup>th</sup> Century Printed Books Dataset”, The British Library YouTube Channel, Accessed August 28, 2019, [https://youtu.be/OGzsDQ\\_Nhyk](https://youtu.be/OGzsDQ_Nhyk).

```

<SP ID="P43_SP00007" HPOS="484" VPOS="177" WIDTH="14"/>
<String ID="P43_ST00009" HPOS="498" VPOS="156" WIDTH="17" HEIGHT="21"
  CONTENT="a" WC="1.00" CC="0"/>
<SP ID="P43_SP00008" HPOS="515" VPOS="177" WIDTH="10"/>
<String ID="P43_ST00010" HPOS="525" VPOS="145" WIDTH="80" HEIGHT="32"
  CONTENT="kind" WC="0.95" CC="0100"/>
<SP ID="P43_SP00009" HPOS="605" VPOS="177" WIDTH="10"/>
<String ID="P43_ST00011" HPOS="615" VPOS="145" WIDTH="39" HEIGHT="32"
  CONTENT="of" WC="1.00" CC="00"/>
<SP ID="P43_SP00010" HPOS="654" VPOS="177" WIDTH="7"/>
<String ID="P43_ST00012" HPOS="661" VPOS="145" WIDTH="157" HEIGHT="43"
  CONTENT="mimicry," WC="0.80" CC="06050102"/>

```

Stripped of location data, this is the phrase “It is indeed a kind of mimicry” on the 43<sup>rd</sup> page in the 1804 edition of Joseph Addison’s *Works*. Without this knowledge of XML parsing, I would have needed to bring someone else on board to provide access to this corpus, and so it is important to account for these types of technical barriers when planning a distant-reading project.

The third methodological consideration I faced was how to divide these full-text manuscripts into time periods for analysis over time. I chose three periods, 1800-09, 1850-59, and 1890-99, as a compromise between feasibility and reliability. On the one hand, it would have been better to consider more time periods so that changes in discourse could be traced to a more precise point in time. On the other hand, time restraints on the scope of this case study limited me to plotting only 300 invocations of civilizing discourses, and for every time period I included, fewer invocations from each time period could be plotted. For example, if I had included 10 time periods, one for each decade of the 19<sup>th</sup> century, I would only have been able to plot 30 invocations each. The overall sample size of 300 invocations for the entire century is already hardly a representative sample, acceptable only because this case study is an exercise in methodological reflection. 30 invocations per 10 years, however, would have been so shallow as to be preposterous. If, however, I had limited the scope to two time periods, the beginning and the end of the century, in order to ensure maximum invocations per time period, I would have lost all ability to plot patterns. Any trend between two time periods is merely a straight line, and a straight line through an entire century is so unnuanced as to be practically meaningless. As such, I chose a compromise by plotting 100 invocations for three time periods.

Once the manuscripts had been divided by time period, the fourth methodological consideration I faced was how to identify 300 invocations of civilizing discourses, and how to annotate them for meaning. Using Python, I identified every mention of ‘civilizing’, ‘civilized’, and ‘civilization’ within each time period, and extracted the paragraphs in which the keywords occurred into XML documents. The word ‘civilizing’ was fairly unproblematic to look for, as it is typically only used in the sense of a civilizing mission. The inclusion of ‘civilized’ and ‘civilization’ was more problematic, because they are used in many contexts besides that of a civilizing mission. By manually reading and annotating these excerpts, however, I was able to ensure that all the excerpts I included were actually invocations of civilizing discourses. Next, I used Python to randomly select 100 excerpts for annotation from each time period. For every excerpt, I manually read the immediately surrounding text within which the keywords occurred to identify the usages of the keywords and, more importantly, the meanings that were applied to them. Identifying meaning and creating categories to store these data points is always a compromise. Amélie Zöllner-Weber, for instance, notes that complex cultural meaning cannot be categorized without losing important nuance, and that individual perspectives lead different people to categorize meaning in different, unpredictable ways.<sup>1</sup> However, as Franco Moretti asserted, “reality is infinitely rich; concepts are abstract, are poor. But it’s precisely this ‘poverty’ that makes it possible to handle them, and therefore to know”.<sup>2</sup> In order to be plotted on a chart and thereby reveals patterns, meanings have to be simplified into categories. I developed two strategies to increase the transparency of my process. First, I did not create an ontology of preconceived categories, but generated new categories as they became necessary to classify an excerpt. By doing so, I tried to sidestep what Øyvind Eide calls the “straightjacket” of ontological categorization, in which the established categories limit what can be plotted.<sup>3</sup> Second, I approached each snippet of text with the same question: what is the author seeming to equate civilization

<sup>1</sup> Amélie Zöllner-Weber, “Ontologies and Logic Reasoning as Tools in Humanities?”, *Digital Humanities Quarterly* 3, no. 4 (2009): §§3-4, <http://digitalhumanities.org:8081/dhq/vol/3/4/000068/000068.html>.

<sup>2</sup> Moretti, “Conjectures on World Literature”, 58.

<sup>3</sup> Øyvind Eide, “Ontologies, Data Modeling, and TEI”, *Journal of the Text Encoding Initiative* 8 (2015): §§39-44, <https://journals.openedition.org/jtei/1191>.

with? For every new answer to this question, I would create a new category of meaning. I will give an example of how this worked. The following is an invocation of civilizing discourse that I excerpted from *Recollections of a Three Years' Residence in China; Including Peregrinations in Spain, Morocco, Egypt, India, Australia, and New Zealand*, published in 1853 by William Tyrone Power. This particular paragraph is a representation of the Maori of New Zealand in the blatantly offensive manner typical of the period.

Their past extending scarcely beyond their own memories; their present a state of wonderful transition; their future one of great doubt yet hope; they afford materials for more interesting study and speculation than those of any other savage race. Here, at least, there is a promise that the child of the soil shall have fair play; that civilization shall take the shape of education and amelioration; that the bent of his naturally good disposition shall be trained in the right direction. We see almost a whole people from the depths of Atheism adopt Christianity in its purest and most uncompromising form.<sup>1</sup>

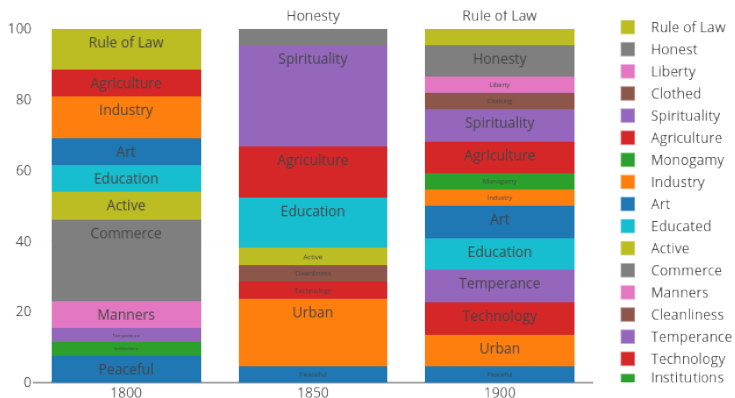
In this excerpt, I identified three categories of meaning. The line, “there is a promise that the child of the soil shall have fair play”, seems to equate civilization with agriculture. The line, “civilization shall take the shape of education and amelioration”, seems (quite obviously) to equate civilization with education. And the line, “we see almost a whole people from the depths of Atheism adopt Christianity in its purest and most uncompromising form”, seems to equate civilization with spirituality. These categories may seem vague, but I had to be careful not to create too many new categories, otherwise, any visualizations created would be overcrowded and would lose the capacity to show trends. For instance, consider two hypothetical snippets of civilizing discourse. The first suggests that to civilize a population, the civilizer must spread belief in Christianity. The second suggests that to civilize a population, the civilizer must end beliefs in magic or practices perceived as superstitions. These are two issues that would, in a traditional research methodology, be addressed separately. In this context, however, if these two snippets are categorized so specifically as to be separate, then each of the remaining 298 snippets would have to be categorized with equal specificity, and we would end with up to 300 unique data

<sup>1</sup> William Tyrone Power, *Recollections of a Three Years' Residence in China; Including Peregrinations in Spain, Morocco, Egypt, India, Australia, and New Zealand* (London: Schulze and Co., 1853), 351.



points that cannot show any change over time. Therefore, the social values I identified had to be categorized using common denominators. Rather than the two examples above being classified as relating to two specific meanings, such as Christianity and magic, they would have been both classified as relating to spirituality.

Once I had applied this method to all 300 excerpts, the final step in my case study was plotting the frequency that each category of meaning appeared over time. My findings are represented in Figure 1 below.



As I explained above, this case study was designed at a small scale in order to explore methodological considerations, and I caution any interpretation of the findings of this distant reading to keep in mind that I did not analyse enough of the corpus to be considered representative. That being said, a brief interpretation of my findings can be illustrative of how patterns revealed by an exploratory distant reading can be used to create new research questions in the history of British civilizing discourses. Two patterns within this graph stand out. First, there are only three values that remain constant throughout the century: agriculture, education, and peace, while many other values fluctuate over time. This is unsurprising, as these are widely accepted by historians as fundamental aspects of the civilizing mission. It is interesting, moreover, that this pattern adheres to Jouni-Matti Kuukkanen’s explanation of conceptual change. Kuukkanen argued that historical concepts are made up of core compo-

nents and marginal components. “The core is something that all instantiations must satisfy in order to be ‘the same concept’ [...] the margin is composed of all the rest of the beliefs that an instantiation of [a concept] might have”.<sup>1</sup> If we consider agriculture, education, and peace to be the core components of British conceptions of civilizing missions, and all the fluctuating categories of meanings to be marginal components, then exploratory distant reading can offer a valuable means of tracing the ebb and flow of marginal components of British conceptions of civilizing missions, and a range of new research questions present themselves. For instance, why does this graph not indicate that spirituality was a core component? Spirituality only appears to dominate in the 1850s, but historians have long attributed civilizing missions with Christian missionary work that dated back to at least the 17<sup>th</sup> century, and as we have seen, John Comaroff specifically locates a religious conception of the civilizing mission. In this case study, it is likely that spirituality is not represented earlier simply because the sample size was not adequately representative, but this is the type of anomaly that exploratory distant reading can highlight.

The second discernible pattern in my findings is a general shift from tangible to intangible categories of meaning. In the 1800s, commerce is the dominant category. Commerce is a tangible value, it is related to immediate material benefits. However, in the 1850s, spirituality and urbanity are the dominant social values. These values are less tangible. Urbanity, the ability to live in a city rather than in the country, is somewhat related to the material world, but spirituality is about internal rather than external issues. In the 1890s, while no value dominates, more new values appear than in either of the previous decades, and the values that do appear have similar undertones. Some of these values, namely, liberty and honesty, are related to philosophical questions about the ways humans should live, while others, such as temperance and monogamy, are subsets of spirituality. The meanings attached to civilizing missions in the 1890s, therefore, appear much more intangible than those of the 1800s, and slightly more than those in the 1850s. This trajectory of civilizing discourses from tangible to intangible could offer a new way to reconcile the debate between the comparative and the discursive histories over how long Britons maintained a belief

<sup>1</sup> Jouni-Matti Kuukkanen, “Making Sense of Conceptual Change”, *History and Theory* 47, no. 3 (2008): 367, <https://browzine.com/articles/14405959>.

in their civilizing missions. The discursive histories argued the belief in a civilizing mission ended around the 1850s, while the comparative histories argue that it continued until around the turn of the twentieth century. If, however, the meanings applied to civilizing missions shifted from tangible to intangible, this could explain how belief in civilizing missions subsided and continued at the same time: it subsided, in that material manifestations lessened, yet it continued in a more personal, philosophical sense. Again, the small sample size used for this case study renders this argument speculative, but these are the types of insights that an exploratory distant reading can provide the study of civilizing discourses.

## Concluding remarks

Given the focus of the previous historiography of 19<sup>th</sup>-century British civilizing discourses on close readings of a small number of texts, it is clear that applying distant reading methods in order to access ‘the great unread’ can add much to this area of study. In this case study, I have shown that using an exploratory framework of distant reading can generate new research questions and help provide nuanced contexts of conceptual. Just as importantly, I have highlighted some of the methodological questions that must be considered when planning a distant reading project of this sort. While this case study was a small scale preliminary exploration of civilizing discourses and could be directly expanded, there are several other directions distant reading could be applied to the subject. This project was focused on the social values that British authors applied to concepts of civilizing missions, but it did not give any consideration as to how popular civilizing missions were as a subject. Historians of Victorian Britain often claim that most Victorians did not care about the empire until the late 19<sup>th</sup> century. In that case, the texts included in this study may represent a minority of published works, or it may indicate that civilizing missions were discussed more often than historians have previously thought. Thus, further analysis could be done on how integrated/popular civilizing discourses were in Victorian writing. It would also be useful to apply this methodology in a more focused way, such as by comparing distant readings of isolated collections of sources. This study included all manuscripts within the British Library’s

full-text corpus, but by applying distant reading to isolated collections individually (administrative correspondence and missionary writings, for example), a comparative discourse analysis could be performed to identify how civilizing discourses differed between disparate social groups. Similarly, this kind of comparative discourse analysis could be performed on sources from before and after events that are commonly accepted as transformational, such as the Indian Rebellion of 1857. This could provide insight into the impact of historical events on the development of British civilizing discourses. All of these directions are not to be taken at the expense of traditional close reading methodologies, as previously stated. As distant reading projects become more popular and more visible, connections between close and distant reading will be made and historians will develop strategies to incorporate both into their research.

## References

- Braudel, Fernand. "History and the Social Sciences". In *On History*, trans. by Sarah Matthews, 25-54. Chicago: University of Chicago Press, 1980.
- Cabrera, Miguel. *Postsocial History: An Introduction*. Translated by Marie McMahon. Toronto: Lexington Books, 2004.
- Capozzi, Mae and Scott Enderle. "A Distant Reading of Empire". Accessed August 28, 2019. <https://readingfromadistance.wordpress.com/>.
- Comaroff, John. "Images of Empire, Contests of Conscience: Models of Colonial Domination in South Africa". In *Tensions of Empire: Colonial Cultures in a Bourgeois World*, edited by Frederick Cooper and Ann Laura Stoler, 163-197. Berkeley: University of California Press, 1997.
- Eide, Øyvind. "Ontologies, Data Modeling, and TEI". *Journal of the Text Encoding Initiative* 8 (2015). <https://journals.openedition.org/jtei/1191>.
- Hall, Catherine. *Civilizing Subjects: Colony and Metropole in the English Imagination, 1830-1867*. Chicago: University of Chicago Press, 2002.
- Jänicke, Stefan, Greta Franzini, Muhammad Faisal Cheema, and Gerik Scheuermann. "On Close and Distant Reading in Digital Humanities: A Survey and Future Challenges". *EuroVis* (2015). <https://doi.org/10.2312/eurovisstar.20151113>.
- Kuukkanen, Jouni-Matti. "Making Sense of Conceptual Change". *History and Theory* 47, no.3 (2008): 351-372. <https://browzine.com/articles/14405959>.
- Laidlaw, Zoë. "Imperial Complicity: Indigenous Dispossession in British History and History Writing". In *Emancipation and the Remaking of the British Imperial World*,

- edited by Catherine Hall, Nicholas Draper, and Keith McClelland, 131-148. Manchester: Manchester University Press, 2015.
- Michel, Jean-Baptiste, Yuan Kui Shen, Aviva Presser Aiden, Adrian Veres, Matthew K. Gray, The Google Books Team, Joseph P. Pickett, Dale Hoiberg, Dan Clancy, Peter Norvig, Jon Orwant, Steven Pinker, Martin A. Nowak and Erez Lieberman Aiden. "Quantitative Analysis of Culture Using Millions of Digitized Books". *Science* 331, no. 6014 (2011): 176-182.
- Moretti, Franco. "Conjectures on World Literature". *New Left Review* 1 (2000): 54-68. <https://tinyurl.com/moretti2000conjwl>.
- Moretti, Franco. "Style, Inc. Reflections on Seven Thousand Titles (British Novels, 1740-1850)". *Critical Inquiry* 36, no. 1 (2009): 134-158.
- Osterhammel, Jürgen. *Europe, the 'West' and the Civilizing Mission*. London: The German Historical Institute London, 2006.
- Pitts, Jennifer. *A Turn to Empire: The Rise of Imperial Liberalism in Britain and France*. Princeton: Princeton University Press, 2005.
- Porter, Andrew. *Religion versus Empire? British Protestant Missionaries and Overseas Expansion, 1700-1914*. Manchester: Manchester University Press, 2004.
- Power, William Tyrone. *Recollections of a Three Years' Residence in China; Including Peregrinations in Spain, Morocco, Egypt, India, Australia, and New Zealand*. London: Schulze and Co., 1853.
- So, Richard Jean. "All Models Are Wrong". *PMLA* 132, no. 3 (2017): 668-673.
- Zöllner-Weber, Amélie. "Ontologies and Logic Reasoning as Tools in Humanities?". *Digital Humanities Quarterly* 3, no.4 (2009). <http://digitalhumanities.org/8081/dhq/vol1/3/4/000068/000068.html>.



*“Imperial shoe, with lace, gold embroidery, beadwork, and knot of gold lace and tassels”, in T. Watson Greig, Ladies’ Dress Shoes of the Nineteenth Century: with Sixty-Three Illustrations, Edinburgh, D. Douglas, 1900. NYPL Digital Collections (<http://digitalcollections.nypl.org/items/510d47da-3b3a-a3d9-e040-e00a18064a99>).*