

Wikipedia and Encyclopaedism: A genre analysis of epistemological values

Steven J. Jankowski

Thesis submitted to Dr Pierre Lévy
in partial fulfillment of the requirements for the degree
of Master of Arts in Communication

University of Ottawa
Ottawa, Ontario, Canada

May 8, 2013

Abstract

This thesis considers how Wikipedia justifies, structures, and legitimizes its production of knowledge. To do so the thesis positions Wikipedia as a site of conflict over the epistemic values between its wiki and encyclopaedic traditions. Through the literature review, the wiki epistemology is argued to be composed of six values: self-identification, collaboration, co-construction, cooperation, trust in the community, and constructionism. While these values are explicit, encyclopaedism's were not found to be equally defined. To fill this gap, the thesis conducts a genre analysis of encyclopaedism. It first identifies the genre through its communicative purposes to create a universal system of total knowledge and to use this system to educate the public. Second, an analysis of recurrent social contexts within Chambers' *Cyclopaedia* (1728), Diderot & d'Alembert's *Encyclopédie* (1751–72), the *Encyclopaedia Britannica* (1771–), and Wikipedia (2001–) finds that the communicative purposes are achieved through the use of five epistemic values: utility, systematic organization, authority, trust in experts, and consistency. Third, a comparison spanning 240 years between Wikipedia and the *Britannica*'s article headings finds that the value of systematic organization structures Wikipedia's articles using seventeenth century categories of knowledge. Having established two sets of values that determine Wikipedia's production of knowledge, the thesis sets the stage for future research to analyze how Wikipedia's epistemology is articulated in its different production spaces. Ultimately, such research may not only describe the shifting values of Wikipedia's epistemology but also explain how knowledge is transformed and produced in the network society.

Keywords: Wikipedia, knowledge, encyclopaedism, genre analysis, network society, epistemology

To Sheila, Anna & Paige.

Contents

I	Introduction	1
II	Literature Review	8
1	Technology	8
1.1	Heidegger	8
1.2	Latour	10
1.3	Castells	13
2	The Network Society	14
2.1	Network	14
2.2	Society	16
2.3	Power & Conflict	19
3	Wikipedia	21
3.1	Wikipedia as wiki	21
3.2	Wikipedia as encyclopaedia	27
3.3	Wikipedia as a site of conflict	28
4	Summary of the Literature Review	30
III	Methodology	32
5	Discourse	32
5.1	Review of discourse analysis	32
5.2	Spaces of discourse	34
5.3	Approaches to discourse	35
6	Genre	36
6.1	Theory	36
6.1.1	Convention	36
6.1.2	Evolution	38
6.1.3	Identity	39
6.2	Examples of genre analysis	42
6.2.1	Bazerman and the evolution of the scientific article	42
6.2.2	Genette's paratext as a threshold to interpretation	46
6.3	Summary	51
7	Genre analysis	52
7.1	Communicative purposes	53
7.2	Social contexts	53
7.3	Textual structures	54
7.3.1	Procedure	55
7.3.2	Selecting texts	56

7.3.3	Acquiring a representative sample	57
7.3.4	Determining the unit of analysis	61
7.3.5	Developing content categories	66
7.3.6	Limitations	67
IV	Analysis	68
8	Section 1: Communicative purpose	68
8.1	Encyclopaedism	68
8.1.1	The creation of a universal system of total knowledge	68
8.1.2	Educating the public through systems of knowledge	69
8.1.3	Generic membership	71
8.1.4	Antecedent influences	71
8.1.5	Emergence during the Eighteenth Century	73
9	Section 2: Social contexts	75
9.1	Corpus of encyclopaedias	75
9.2	Chambers' <i>Cyclopedia</i> , 1728 CE	76
9.2.1	Lexicon	76
9.2.2	Connected literature	78
9.2.3	Intended audiences	79
9.2.4	Author	80
9.3	Diderot and d'Alembert's <i>Encyclopédie</i> , 1751–72 CE	80
9.3.1	Lexicon	80
9.3.2	Connected Literature	82
9.3.3	Intended Audiences	83
9.3.4	Authors	83
9.4	The <i>Encyclopaedia Britannica</i> , 1771– CE	85
9.4.1	Lexicon	85
9.4.2	Connected Literature	90
9.4.3	Intended Audiences	92
9.4.4	Authors	95
9.5	<i>Wikipedia</i> , 2001– CE	97
9.5.1	Lexicon	98
9.5.2	Connected Literature	104
9.5.3	Intended Audiences	106
9.5.4	Authors	108
9.6	Synthesis of social contexts	115
9.6.1	Overview of contexts	116
9.6.2	Discriminative strategies and epistemic values	122
9.6.3	Conclusions	125
10	Section 3: Textual structures	126
10.1	Results	128
10.1.1	Number of headings	128
10.1.2	Total heads per article per encyclopaedia	128
10.1.3	Total heads per edition per encyclopaedia	129

10.1.4	Number and usage of heads per article per encyclopaedia	131
10.1.5	Percentage of ranks used per edition per encyclopaedia	132
10.2	Common and unique heading words	133
10.3	Most influential heading words	134
10.3.1	Most frequent words common to both encyclopaedias	134
10.3.2	The most frequently used words that are common to both encyclopaedias	134
10.3.3	Most frequent words unique to each encyclopaedia	135
10.3.4	Multiple frequencies	135
10.3.5	Comparison of most influential common and unique words	138
10.4	Conclusions	138
V	Discussion	141
10.5	Theoretical and Methodological Limitations	143
10.6	Summary of Contributions	144
10.7	Future Research	146
10.8	Conclusion	147
VI	End Matter	149
11	Bibliography	149
12	Appendix	163

Part I

Introduction

Since 2012, between ten and fifteen per cent of the daily Internet traffic (Alexa, 2013) has passed through Wikipedia, the sixth most popular web site on the Internet (Alexa, 2012a). This fact is reflected by how Wikipedia is often one of the first places people visit on the Web to learn about unfamiliar knowledge. For example, on April 19, 2013, a number of news sources reported that the Boston Marathon Bombers of April 15th were originally from a region near Chechnya (Washingtonpost.com, 2013). In the space of that same day, a staggering 1.5 million visits were recorded on Wikipedia's "Chechnya" article, almost seven times more than combined visits the article received in the first quarter of 2013.¹ What this surge in traffic illustrates is that Wikipedia is not only a source of quick reference or a tool to satiate idle curiosity. It is also a significant source of knowledge we use to understand important events in our lives.

Underlining the current thesis is a concern for understanding this significant source of information by understanding how it shapes and produces its knowledge. Descriptions of Wikipedia often allude to the fact that its content has been produced with the open, flexible, and transparent characteristics of its wiki infrastructure. This rather remarkable database allows an unprecedented amount of information to be produced through the participation and collaboration of millions of people. At the same time, it cannot be overlooked that Wikipedia also belongs to a tradition of encyclopaedias, a genre that has its own particular forms of production. If both processes of production are at work, what is the resulting shape of knowledge that it produces? How is it structured? What makes it legitimate? Further to the point, what limits are placed on Wikipedia's knowledge by way of its production?

Borrowing from the language of Alex Korzybski, an encyclopaedia provides users a map of the territory of knowledge. While it is true that a "map is not the territory it represents" (Korzybski, 1958, p.58), the map becomes "a factor in the reality it explicates" (Lévy, 2011, p.98). In this sense, an encyclopedia outlines which epistemological roads are frequently

¹The following data of visits to the Wikipedia article "Chechnya" was taken from <http://stats.grok.se/en/201301/Chechnya>; 201302/Chechnya, 201303/Chechnya, and 201304/Chechnya. Visits for January 2013: 69,441; visits for February 2013: 73,206; visits for March 2013: 73,141. The total number of visits for the three months combined: 215,788. Visits on April 19, 2013: 1,508,827.

travelled and which parts of town one should avoid. In other words, an encyclopaedia as popular as Wikipedia plays a part in defining our reality, reinforcing certain knowledge and excluding others. The purpose of this thesis is to reveal this epistemological roadmap, to seek out the limits of Wikipedia's knowledge, and to understand the particulars of how it mediates the world through its pages. While the focus of this thesis is squarely centred on Wikipedia's production of knowledge, it also addresses some broader questions of how knowledge is articulated through the intersections of technology and society.

According to Manuel Castells (1996), society in the late twentieth-century was transformed by advancements made in communication technologies, the globalization of capital, and the increased need for individual freedom and open communication. He defines the result of this transformation as the "network society", a society that is determined by the flexibility, scalability, and survivability of all forms of networks. Because of the convergence of these three processes, a "structural transformation" of society can be witnessed in the new relationships of production, power, and experience that permeate throughout our social lives (Castells, 2000, p.376). A number of scholars, including Yochai Benkler (2006) and Pierre Lévy (2011), argue that such changes are also transforming what knowledge is and how it is produced. In all cases, these scholars point to popular mass collaborative projects like Wikipedia, the user-generated online encyclopaedia, as models of this transformative process. In opposition to this stance are scholars like Featherstone & Venn (2006) who argue that instead of revolutionizing the qualities of knowledge, Wikipedia actually reproduces traditional and hierarchical knowledge structures, structures that run counter to the values of the network society. Given these differences of opinion, how can we understand Wikipedia if it is both "wiki" and "-pedia"? The current thesis argues that these discrepancies can be understood by observing how each side of Wikipedia is guided by epistemic values that seek the communication power to produce knowledge and make meaning. In this view, the contrasting definitions of Wikipedia can be considered manifestations of the process of knowledge transformation at the beginning of the twenty-first century.

In order to ground this examination of the network society, the thesis begins by discussing what technology is and how it relates to society. Contrasting Martin Heidegger to Bruno Latour, the discussion points to the fact that technology mediates social reality, and as such, plays an undeniable role in defining how society functions. By establishing this perspective, Castells'

explanations of why the network society has such transformative power become all the more clearer. Having built a philosophical framework to support Castells, the chapter concerning the network society explains what networks are and how these networks have transformed traditional social roles. Following these descriptions, the rest of the chapter is dedicated to explaining how the network society functions through communication power and is shaped by conflicts between different value systems.

Given the fact that the contours of the network society are defined by conflicting values within networks, the thesis considers that knowledge within the network society has been transformed by this same process. Enter Wikipedia, the often lauded model project of the Internet whose very existence is based on the negotiation of conflicting meanings. Castells himself (as well as Lévy (2011), Benkler (2006), Tapscott (2006) and Shirky (2010)) points to Wikipedia as a symbol of the kind of potential the network society can create. In Tapscott's *Wikinomics* (2006), he defines "wiki", the software package that facilitates mass collaboration projects, as "a metaphor for a new era of collaboration and participation" (p.18). Among mass collaboration, Wikipedia has been attributed such status because it embodies features such as free content (both legally and financially), multiple languages, it is open to anyone² to edit, there is very little vertical hierarchy, users are autonomous, and finally its infrastructure is built around community building and discussion. It is through these characteristics that Wikipedia offers unprecedented freedoms in the kind of knowledge that is collected and talked about by Wikipedians. Following Tapscott's lead, this thesis will hence use the term "wiki" to describe this package of characteristics.

The previous scholars outline that Wikipedia manifests the values of the network society. Additionally, Yochai Benkler writes that it is through the decentralized distribution of work that has lead Wikipedia to produce a "radically new form of encyclopedia writing" (2006, p.70). However, the fact that Wikipedia is an encyclopedia, although obvious, is rarely explained and its implications are even less understood. As Featherstone & Venn point out, even if it utilizes a relatively new medium, Wikipedia continues to follow "traditional disciplinary divisions" and is "low on inter-disciplinarity and innovation" (2006, p.10). In alignment with these statements is the study conducted by Emigh & Herring (2005). This study found that when comparing

²There is a caveat on who this "anyone" is, which will be explored in the chapter on Wikipedia.

articles from the Online Columbia, Everything2, and Wikipedia the writing on the Columbia and Wikipedia were not “significantly different from one another. Statistically speaking, the language of the Wikipedia entries is as formal as that in the traditional print encyclopedia” (Emigh & Herring, 2005, p.5). They continue by describing that both the “Columbia Encyclopedia and Wikipedia entries are systematic, standardized, and narrow in scope” (Emigh & Herring, 2005, p.9). What these scholars argue is that in stark contradiction to the characteristics evident in wikis and the network society at large, Wikipedia has not transformed the shape of knowledge.

If the question of this thesis is to understand how knowledge is transformed in the network society, the preceding scholarship poses an apparent problem for using Wikipedia to understand this process, as Featherstone & Venn argue, Wikipedia has done no such thing. In fact, Emigh & Herring’s 2005 article presents this issue as a paradox: “How is it that the wide-open participation structure of a wiki can reproduce traditional print norms?” (2005, p.9). What these contradictions in scholarship show is that there is a question of the degree that Wikipedia’s production of knowledge is the result of being a wiki or due to being an encyclopaedia. This issue raised by the contrasting scholarship becomes the pivot of the current thesis. As such, it forms the basis of the research question of the thesis:

What is the epistemology of Wikipedia?

To answer this question, the thesis returns to Castells’ explanation that the network society is defined by the conflicts between networks of values. Because cultural meaning is made through the meshing process of adjacent networks, the thesis establishes that the epistemology of Wikipedia is the result of a network conflict, one between the epistemic values of wiki and the epistemic values of encyclopaedism. It is the purpose of the thesis to make these sets of epistemic values explicit. In doing so, it establishes the ground upon which epistemic conflicts on Wikipedia occur, ultimately describing a case study in how knowledge is transformed in the network society.

However, to do so, both sets of values need to be outlined. In the case of wikis, the current study supports Ruth & Houghton’s (2009) outline of the epistemology of wikis with a number of additional studies to establish wiki’s epistemic values: self-identification, collaboration, co-construction, cooperation, trust in the community, and constructionism. In the case of encyclopaedism’s epistemic values, the researcher has not been able to locate a comparable

articulation of encyclopaedism's epistemic values. This lack of comparability makes the establishment of Wikipedia's epistemology a difficult task. Therefore the bulk of the current thesis is dedicated to the following questions:

What is the genre of encyclopaedism? and; What is the epistemology of encyclopaedism?

To build a methodological framework to understand encyclopaedism, the current study again returns to Castells' explanation that the conflicts between networks of values occur within discourse. If this holds true, then the values of encyclopaedism should be found within their discursive practices and productions. As such, the methodology starts by broadly defining discourse by relying on the scholar Vijay Bhatia (1993 & 2004). The chapter on discourse uses Bhatia's description of the four spaces of discourse and four approaches to discourse analysis to develop a methodological framework. Of all the approaches to discourse analysis, Bhatia advocates for a generic perspective as it takes into account each of the four spaces of discourse. As such, a discourse analysis through genre reveals how the textual production and social practices of a genre produce meaning.

Having established that the main methodology of the thesis is framed by genre, the chapter on genre describes how they work and how they can be identified. Relying on Bhatia (1993), Miller (1984), and Askehave & Swales' (2001) definitions of genre, they state that a genre can be identified by its "communicative purposes". Additionally, Bhatia explains that such purposes are enacted through discriminative strategies. Therefore, if a discriminative strategy is changed, then the communicative purpose is also changed, leading to a new genre. Armed with these theories of genre, the thesis reviews two forms of genre analysis, one that concentrates on social contexts by Charles Bazerman (1988) and another that concentrates on textual structure by Gérard Genette (1997a). In Charles Bazerman's investigation of the emergence of the scientific article (1988) he uses a variety of studies to reveal the epistemic decisions of the discursive community of scientists. One such analysis explains that a genre's epistemology can be discovered through the recurrent contexts of lexicon, connected literature, intended audience, and authors. Additionally, in a textual driven analysis, he also discovers that these same epistemological choices are present in the way that scientific articles are structured through headings. This perspective of headings is similar to the function that Genette describes of paratexts (1997). He explains that textual devices like headings act as "thresholds of interpretation" (Genette, 1997), places where the

intentions of the author (the communicative purposes) are liminally established.

With the methodological framework firmly established, the thesis identifies the genre of encyclopaedism through its communicative purposes, reveals its epistemic values by investigating how it addresses the four recurrent contexts, and validates these epistemic values by observing how they are manifested in the textual structure of encyclopaedic texts. Therefore, the first section of the analysis sets out to define encyclopaedism as a genre. To do so it relies on work of historians Doody (2009), Yeo (2007) and Burke (2000) to aid in defining the communicative purposes of encyclopaedism. It finds that there are two communicative purposes that encyclopaedias set as their goals. First, the creation of a universal system of total knowledge and second, educating the public through systems of knowledge. Guided by these purposes, the thesis delimits the generic membership, explains which older genres influenced its formation, and a brief description of the social contexts that it emerged from.

The second part of the analysis follows Bazerman's methodology of four contexts through four canonical encyclopedias: Chambers' *Cyclopaedia*, Diderot & d'Alembert's *Encyclopédie*, the *Encyclopaedia Britannica*, and Wikipedia. After describing in detail how each encyclopaedia embodies the four contexts, the analysis synthesizes the differences of these contexts as the evolution of the genre. Additionally, those aspects that appear to be consistent between the four encyclopaedias are considered to be the discriminative strategies of the genre and thus reveal its epistemic values. As such, the study describes that encyclopaedism has five main epistemic values: utility, systematic organization, authority, trust in experts, and consistency. While these values may be hardly surprising, the details that make up these values tend to differ from most conceptions of encyclopaedias. For instance, while encyclopaedias are usually commented to be organized by topic and alphabetization, these are but two of several systematic organizations that encyclopaedias use. Additionally, the study of social contexts revealed that the value of trust in experts is not solely reliant on expert contributors as the question posed by Emigh & Herring assumes, but on expert sources. Thus the current thesis provides an in-depth and nuanced analysis of the discriminative strategies and therefore the epistemic values of encyclopaedism.

The third and final part of the analysis is used to observe how these epistemic values are manifested in the paratextual structure of both the *Encyclopaedia Britannica* and Wikipedia's

article heading hierarchy. The results show that the epistemic values of consistency and systematic organization weigh heavily on editors of both the *Britannica* and Wikipedia. With the chapter on genre in mind, this is hardly surprising since in order for an encyclopaedia to be an encyclopaedia it must adhere to its discriminative strategies. As Section 2 of the analysis describes, consistency and systematic organization are discriminating qualities and are necessary for Wikipedia to be an encyclopaedia.

The discussion section of the thesis elaborates on the findings of this thesis by using Emigh & Herring's (2005) question to illustrate how using discriminative strategies can provide a valuable framework for understanding what Wikipedia is and how it functions. The discussion concludes by indicating a number of epistemic conflicts that future research can now explore given that the two epistemic value sets have been firmly established.

To summarize, this thesis attempts to answer the question of how knowledge is transformed in the network society. To do so, it looks to Wikipedia as an example where knowledge is being transformed by the conflicts over the epistemic values of two networks, encyclopaedism and wiki. However, while the epistemic values of wikis have been succinctly defined, encyclopaedism's epistemic values are mired in historical inconsistencies of genre definitions. As such, the majority of the analysis is dedicated to answering "What is the genre of encyclopaedism". Resulting from this question is the establishment of the two epistemic sets of values that allows future research to answer the questions of "What is the epistemology of Wikipedia?" and how does this epistemology represent the transformation of knowledge within the network society.

Part II

Literature Review

1 Technology

1.1 Heidegger

If the purpose of current thesis is to understand how transformations within the network society also transform the production of knowledge, the concept of the network society needs to be unravelled. To do so, one must question what is the nature of the relationship between technology and society. To answer this question we must ask yet another question that goes to the heart of the matter; “What is technology?” In answering this question, Heidegger (1977) warns in his influential chapter “The question concerning technology”, a preoccupation with the *things* of technology creates a barrier to being fully free from our relationship to technology. He adds that it stops us from being masters of technology. In Heidegger’s words we are “unfree and chained to technology, whether we passionately affirm or deny it [...] we are delivered over to it in the worst possible way when we regard it as something neutral” (Heidegger, 1977, p.287). In other words, by concentrating on just the material and shape of technology, we fail to recognize the greater complexity of *what* exactly, technology is. Instead we are trapped, locked into a relationship with technology that constitutes humans as a stable resource ready to be called upon at the whim of technology.

Heidegger explains this through a spiralling ontology using Greek, German, and Latin, where he marks out a philosophical path that steadfastly explains that the essence of technology is to reveal truth. He arrives at this conclusion by first paring from technology two definitions, one that defines technology as “a means to an end” while the other one defines it as “a human activity” (Heidegger, 1977, p.287). For Heidegger, the instrumental and the anthropomorphic “definitions of technology belong together” (1977, p.287). Following the instrumental line, he describes in succeeding steps how technology is a way of revealing truth. Given that the instrumental definition of technology states that it is a means to an end, he states that “ends” and “means” belong to the realm of “causality”. From causality, anything that causes something to happen is an effect, or an “occasioning” (Heidegger, 1977, p.289). Such “occasioning” is the

setting-in-motion of something, or a bringing-forth. Bringing-forth leads to revealing, and so, in revealing we find truth. Through this lineage Heidegger establishes that technology “is a mode of revealing. Technology comes to presence in the realm where revealing and unconcealment take place, where [...] truth, happens” (1977, p.295). This brief summary of his inquiry leads to Heidegger’s issue with technology.

The problem that Heidegger identifies in this chain of revealing is that not all technology reveals in the same way. On one hand, there is “poesis” which provides a revealing through opening (Heidegger, 1997, p.293). It includes the bursting into the open of something in and of itself (like the blooming of a flower), and “the bursting open belonging to bringing-forth, not in itself, but in another [as] in the craftsman or artist” (Heidegger, 1997, p.293). Alternatively there is revealing through enframing, and it is this particular revealing that creates a barrier to mastering technology.

Heidegger states that the “revealing that rules in modern technology is a challenging, which puts to nature the unreasonable demand that it supply energy which can be extracted and stored as such” (1977, p.296). Heidegger explains that modern science, the precursor to modern technology “pursues and entraps nature as a calculable coherence of forces” and as such, when this form of challenging is in place, nature cannot be revealed in any other manner (1977, p.302). As much as humans are part of the modern act of ordering nature, we are not immune to being subsumed by that exact same process. In fact we too are swept up into this ordering which Heidegger calls “Gestell”, or enframing (1977, p.301). Because modern technology reveals through enframing, both humans and nature are set to be revealed in a particular way. He provides the example of the Rhine, a river whose essence can only be derived from the fact that a power station derives hydroelectric power from it (Heidegger, 1977, p.297). Taken a step further, and implied by Heidegger, the men and women who work in the power station are themselves, a human resource or a labour force; collected, put to work, and dismissed in the same way that the river is. In both cases modern technology has turned both river and humans into standing-reserves, stockpiles in waiting, always ready to be used by technology.

To summarize, Heidegger envisions that there is an essential and a priori truth, one that eternally exists and it is through technology that we gain access to it. However, because of the nature of modern technology, truth has been revealed through enframing. Thus the reality we

have access to is constituted only in terms of the frame. We could change this situation if we were able to wrest away revealing from enframing and return it to a bringing-forth that is an opening. Therefore, Heidegger argues that we should return to the revealing that craftsmen and artists are capable of. This switch would allow us to take hold of and master technology once more, aligning ourselves to the open revealing of truth.

1.2 Latour

For Bruno Latour, much of what Heidegger says about technology is hogwash. His major point of contention with Heidegger's position is that Heidegger describes technology as something that lies outside of human action, that it is "insuperable, omnipresent, superior", and "a monster born in our midst which has already devoured its unwitting midwives" (Latour, 1999, p.176). Furthermore, Latour captures Heidegger's sentiment when he recapitulates that to "become moral and human once again, it seems we must always tear ourselves away from instrumentality, reaffirm the sovereignty of ends, rediscover Being; in short, we must bind back the hound of technology to its cage" (2002, p.247). Latour argues that the contrary is true, that we "never tame technologies, not because we lack sufficiently powerful masters, not because technologies, once they have become autonomous, function according to their own impulse, not because, as Heidegger claims, they are the forgetting of Being in the form of mastery, but because they are a true form of mediation" (Latour & Venn, 2002, p.250). This "true form of mediation" that Latour mentions contends that instead of being mobilized *by* technology, we act *with* technology, that it mediates all of our relationships, whether it is with other humans or with the world around us.

In contrast to Heidegger, Latour states that in "artifacts and technologies we do not find the efficiency and stubbornness of matter, imprinting chains of cause and effect onto malleable humans" (1999, p.190). Instead, Latour argues that this kind of separation between instrumentality and human action has led us down the wrong path to understanding the relationship between society and technology. As the editors of *Technology and society: building our sociotechnical future* describe, Latour argues that technologies "play such an important role in mediating human relationships [...] that we cannot understand how societies work without an understanding of how technologies shape our everyday lives" (1992, p.151). There lies the importance of this

reconfiguration of what is technology: “If we abandon the divide between material infrastructure on the one hand and social superstructure on the other, a much larger dose of relativism is possible” (Latour, 1991, p.129). The relativism that he speaks of is reconstituting Heidegger’s instruments as pivotal actors *within* society, not as a force that resides and acts from without.

To elaborate, Latour describes how the most familiar actor in social theory is of course the human. Latour explains that as most social theories myopically concentrate on the agency of humans, they willingly exclude a substantial aspect of what makes society durable. Latour’s Actor Network Theory (ANT) includes these “missing masses”, or what he calls “non-humans”, actors within society that are not human but retain just as much agency (Latour, 1999). Through a series of technical mediations³ humans and non-humans act together to create new social situations that sustain the structure of society. He states that “whenever we discover a stable social relation, it is the introduction of some non-humans that accounts for this relative durability” (Latour, 1991, p.110). In this regard, technology is pivotal for creating meaningful relationships with the world.

Latour’s classic example of ANT at work is the slogan used by the NRA “guns don’t kill people, people kill people” (1999, p.176). From the NRA’s perspective, guns are a neutral object that are wholly mediated by the goals of a human subject. Whether a person was killed by a gun or not is left up to whether the person holding the gun is either “good” or “bad”. The traditional opposition to this stance is that when a “good person” picks up a gun the “virtue of *material* components” will now make the person dangerous (which is not unlike Heidegger’s view of technology). Latour describes how both positions try to place responsibility by making one side of the equation neutral or powerless and the other powerful. The truth is that without the human, the gun could not shoot someone, but without the gun, the human could not kill using a gun. Therefore Latour rephrases the relationships as being two actors, a human and a non-human, and in acting together they “interfere” or detour one another’s goals, thus creating a third actant, the gun-human who has a completely different set of abilities that comes from the original actors (Latour, 1999, p.178). By describing our relationship to technology in this way, Latour avoids the object-subject and neutral-active categories that do nothing to explain how a situation

³Latour describes that there are four meanings of technical mediation: interference or detour, composition, folding time and space, crossing the boundary between signs and things (Latour, 1999, pp.178–185). Each of these mediations describe how actors come together to form new networks.

actually comes to be (Latour, 1999, p.194).⁴

As seen by the combination of gun and human resulting in a third actor, the gun-human, we see an example of “blackboxing”. This term is used by Latour to describe how technical action folds “Time, space and the type of actants” (Latour & Venn, 2002, p.248). To make this clearer, let us return to the gun in the first example. The gun is not *just* a gun, but is the result of a variety of non-human actions and human actions. Folded into the gun is the ability to kill at a distance, and so spears, gun powder, springs, arrow heads, loading mechanisms, iron-casting, metallurgy, are all present to achieve this goal. But so to are the engineers, inventors, military institutions, economies, contracts and laws that combined (through an endless series of interferences) to result in the gun being held by our exemplar gun-human. Blackboxing folds all of these actors into a single actor, so that all we can see is an opaque box in front of us, and not the near-infinite chain of relationships and actants that are present. It is through these chains of action folded into blackboxes that Latour announces that instead of the monster of modern technology that Heidegger vilifies, technology is the process with which we commune with the world around us. It is through technology that society is made real.

While it is obvious that Heidegger and Latour are diametrically opposed in many respects, they each hold a similarity to the other, they both position technology as central to defining reality. In the case of Heidegger, technology defines the degree of verisimilitude between the truth and the revealed truth. By his definition, the current reality is due to modern technology’s need to enframe everything. For Latour, each connection between actors adds to the mesh of reality.⁵ Thus, blackboxes that fold the most chains of connections are in the position to detour the conditions of other actions, and thus define the contours of reality.

⁴Approaching social theory from this angle redirects questions whose answers can only be arbitrary categories (such as “‘is this social’, ‘is this technical or scientific’, or asking ‘are these techniques’”) and turns them toward questions that unfold our technological relationships: “has a human replaced a non-human? has a non-human replaced a human? has the competence of this actor been modified? has this actor human or non-human been replaced by another one? has this chain of association been extended or modified?” (Latour, 1991, p.110).

⁵Latour calls these meshes of human and non-human networks collectives (Latour, 1999, p.198).

1.3 Castells

At the beginning of the twenty-first century, we are well aware of one such blackbox, the network society.⁶ As Castells explains, the current epoch is not just a result of the technical infrastructure incurred by the Internet, but also “the needs of the economy for management flexibility and for the globalization of capital, production, and trade” and “the demands of society in which the values of individual freedom and open communication became paramount” (2001, p.2). These three “independent processes” have been folded into the network society and provide “the common frame of reference for symbolic processing from all sources and all messages” and “thus of our experience as communicating beings” (Castells, 2000, p.693). This description of the network society by Castells is in keeping with Latour’s ideas about technology. Castells makes the point that usual definitions of technology revolve around “the use of scientific knowledge to set procedures for performance in a reproducible manner” and that it is “linked to the conditions of scientific discovery, technological innovation, and application and diffusion in society at large” (2004, p.8). Aligning himself with Latour, Castells points out that this perception of technology wrongly makes a categorical distinction between things that are *social* and things that are *technological*. Quoting Claude Fischer, he states that technology is not “an exogenous factor affecting society” but is “a socially embedded process” (Castells, 2000, p.693) or even more “it is society” (Castells, 2006, p.3). Further to the point, if technology is embedded in society, it stands to reason that society “shapes technology according to the needs, values, and interests of people who use the technology” (Castells, 2006, p.3). This is important for Castells because when we adjust current communications technologies for our needs and interests “the actual modifications introduced in the technology are communicated back to the whole world, in real time” (Castells, 2001, p.28). Thus, for Castells, technology “is a fundamental dimension of social structure and social change” (2004, p.8). Through these definitions of the network society

⁶Interesting to note, that in Tim Berners-Lee’s proposal (1989) for what would become the World Wide Web, he similarly makes it explicit that “linked information systems” are useful because the circles [nodes] and arrows [edges] can stand for anything” (Berners-Lee, 1989). Furthermore, like Latour, he does not make a distinction between humans and non-humans as active participants in the network. He states that a node “represents or describes one particular person or object. Examples of nodes can be: People; Software modules; Groups of people; Projects; Concepts; Documents; Types of hardware; Specific hardware objects” (Berners-Lee, 1989). Here, at the very gestation of the network society, the logic of the network was already at work, working out how to describe and explain the relationships between humans and non-humans.

we see that Castells has much in common with Latour's position on technology.⁷ As well, the network society is responsible for determining a great variety of connections between ourselves and non-humans, and amongst non-humans.

In summary, this section has outlined two philosophical positions about technology. From Heidegger we come to understand technology as a way of revealing truth, but in its modern condition, technology controls us through enframing, which is detrimental to our Being as it remakes us as a standing-reserve. Latour conversely argues that it is *through* our relationships amongst humans and non-humans (this includes Heidegger's instruments) that we establish reality; that anything that we need to know or to act is located within the technical mediations that are the mesh of society. We have seen that Castells follows a similar understanding of technology as Latour and in doing so, sets up a condition that is found in all three authors, that technology is central to the constitution of social relationships (society itself). Now that the nature of this relationship is established, the following section describes in greater detail the particulars of the network society: what exactly is a network, how it functions and how it has become the structure that defines our current society.

2 The Network Society

2.1 Network

To understand the network society, it stands to reason that one must understand networks. In the last fifteen years, there have been numerous works that emphasize the network as a pivotal structure to understand the relationship between society and technology. Manuel Castells' *The Network Society* (2004) and Yochai Benkler's *The Wealth of Networks* (2006), are just a couple titles that position the network as the defining structure of the current society.

The definition of the network can be considered in a number of ways. A professor of media

⁷Although similar, there is an ontological difference between how Castells and Latour relate society to technology. Castells is more closely associated with the social constructionist school of thought as it describes that while there is a reciprocal relationship between society and technology, this relationship is still based on the difference between subjects and objects. Latour argues that this configuration is artificial and continues the myth of a separation between humans and the world they live in (Latour, 1999, Chapter 1). For him, social constructivism remains an incorrect model for understanding the relationship between technology and society. He explains that those "advocating the actor network approach agree with the social constructivist claim that sociotechnical systems are developed through negotiations between people, institutions, and organizations. But they make the additional interesting argument that artifacts are part of these negotiations as well" (Latour, 1992, p.151).

analysis, Joost van Loon, has written a comprehensive summary of the network, stating that it is “a device for organizing and conceptualizing non-linear complexity” (2006, p.307). He goes on to describe how the “multiplicity of traces” defy “the traditional linear structures of narrative, chronology, and genealogy” (van Loon, 2006, p.307). Its structure is composed of three elements; links, nodes, and mesh (van Loon, 2006, p.307). He explains that links are the feature that bind the network together; nodes are the points where links congregate; and the mesh is the form that the connected nodes and links make when seen as a whole.

The network is also sometimes described in terms of graph theory. From this perspective, a network is organized by the various levels of connectivity between edges (links) and nodes and their overall topography (or mesh) (Galloway, 2007, p.32). Galloway points out that the topography of the network is dictated by the amount of equality between nodes. He derives this definition directly from Paul Baran, who wrote a report on the US military telecommunications system in 1964. The report explains the best way to construct a communications network that is resilient to nuclear bomb attacks on American soil. From this report Baran (1964) outlines three forms of networks. The most familiar arrangement is the centralized network, one with a powerful node that is central to all connections. The concept of the hierarchy or tree is a prime example of this kind of network, with each branch being subservient to the main branch or trunk. A second form is the decentralized network, which is similar to the centralized form except that more than one node mediates the activity of the network. The final form is the distributed network where every node is equal to all others, extinguishing the need for a center to mediate connections. The distributed arrangement, according to Baran, is the optimal network that is robust and redundant enough to withstand a nuclear attack.

In 1980, French philosophers Gilles Deleuze and Felix Guattari wrote an influential philosophical treatise, *A Thousand Plateaus* (1980). In this work, they describe the metaphor of the rhizome, a type of root system that “connects any point to any other point”, which is different from the traditional hierarchical root system of trees. If one substitutes the word “rhizome” for “distributed network”, the chapter on the rhizome can easily be understood in terms of the Internet. The two philosophers state that the rhizome “does not have a structure of set points and positions” but “lines of segmentarity” that undergo constant metamorphosis (Deleuze & Guattari, 1980, p.9). Characteristically, the constantly moving structure of the rhizome “operates

by variation, expansion, conquest, capture, offshoots” (Deleuze & Guattari, 1980, p.23). These sorts of motions are the same kind that exist in today’s distributed networks.⁸

2.2 Society

With these definitions in mind, Barney explains that “the spirit of our age is the spirit of the network” and that “the constitutive principles of networks” animate all aspects of our lives (2004, p.2). As such, Castells explains that our current society is defined as a network society, one “whose social structure is made of networks powered by micro-electronics-based information and communication technologies” (2004, p.3). He defines social structure as “the organizational arrangements of humans [which are] expressed in meaningful communication coded by culture” (Castells, 2004, p.3). He later explains that the Internet, the current popular form of communication, is based around three principles: “a decentralized network structure; distributed computing power throughout the nodes of the network; and redundancy of functions in the network to minimize the risk of disconnection” (2001, p.17). Taken individually, he describes that first, our relationships between one another are linked by digital information via communication technologies. The information is in turn, coded by the culture that uses it. These links between cultural codes, information, technology, and people create a topography (mesh/graph) that can be described as a decentralized network society, one with many hubs of activity and importance.⁹ To this point, Castells is describing a mixture of network types interacting with one another.

What makes this form of society and the technology of the Internet different from others is that it has a great degree of flexibility or elasticity. As Castells states, the “elasticity of the Internet makes it particularly susceptible to intensifying the contradictory trends present in our world. Neither utopia nor dystopia, the Internet is the expression of ourselves—through a specific code of communication, which we must understand if we want to change our reality” (2001, p.6). Such elasticity is seen in the simultaneous communication models at work. Unlike

⁸In addition to the rhizome, the authors distinguish between two other structures: the first two being the “centered” (read:centralized) and “polycentric” (read:decentralized) “systems with hierarchical modes of communication and preestablished paths” (Deleuze & Guattari, 1980, p.23).

⁹To clarify what Castells is saying, van Loon describes Castells’ society networks as “both territorialized in particular centres of economic activity and trade, but also deterritorialized in global flows of capital, goods, information, symbols and people”, calling attention that our society is not purely of one type of network, but a mixture of decentralized and distributed (van Loon, 2005, p.309).

the television broadcast model of one-to-many, the Internet includes combinations of one-to-one, many-to-many, many-to-one, and from anyone(read:many)-to-anywhere (Castells, 2001, p.2). Wellman *et al.* describe this process as a part of the communities and societies that “have been changing towards networked societies where boundaries are more permeable, interactions are with diverse others, linkages switch between multiple networks, and hierarchies are flatter and more recursive” (2003, sec. The Rise of Networked Individualism, para. 4). What this affords the network society is an adaptability to different scenarios and contexts that are no longer limited by the technology that mediates it.

What Castells considers paramount to the elasticity of the Internet is the ability to incorporate changes that happen in real-time (Castells, 2001, p.28). This explains the speed of exchange on the Internet. This self-reflexive action is termed “feedback”, a mechanism that is fundamental to the growth of the communication infrastructure of the network society.¹⁰ This is due to the use of electronic media that allows near-synchronous and asynchronous communications. For Castells these reconfigurations of time has created “the systematic perturbation in the sequential order of phenomena” which he calls timeless time (1997, p.183). Thus the network society and the communication technologies it employs are flexible and elastic, even in terms of its temporal existence.

Lévy’s recent work *The Semantic Sphere* (2011) points out that because “the digital medium [is] fundamentally participatory, welcoming of diversity and impossible to shut down [and] a medium for collective creativity”, we must learn to take advantage of its potential (p.21). It is under these conditions of an elastic and digitally networked social fabric that Lévy prescribes the ability for us to take full advantage of the collective intelligence of humanity. In Lévy’s words, collective intelligence is the conjunction of “individual intelligences that are very real but whose activity only becomes meaningful in interdependence with the thinking societies” and the “symbolic systems” that are shared between them (2011, p.196). In other words, collective intelligences are the “cognitive capacities of a society, a community or a collection of individuals” (Lévy, 2011, p.90). Here, cognitive activities include clarification, vision (envisioning), problem solving, hypertextual consistency (interconnectedness), trust, and reciprocal unity (Lévy, 2011,

¹⁰The concept of feedback has been an important aspect of cybernetics since Norbert Wiener described it in 1954. He defined feedback as “the property of being able to adjust future conduct by past performance” (Wiener, 1954, p.33). For him, feedback allowed a system to evade entropy, or the degradation of a signal.

p. 24). All of these activities are already occurring in varying degrees on the Web. For instance he provides the striking example of Wikipedia which is exemplar in illustrating “the power of collective intelligence emerging from a civilized creative conversation [and] in which authors, readers and editors exchange roles to further the dissemination of knowledge” (Lévy, 2011, p.91). In order for collective intelligence to be achieved, and he stresses this point, it can only be done so through creative conversation, by “coordinating unique elements and facilitating dialogue, and not by levelling differences or silencing dissenters” (Lévy, 2011, p.92). It is through these terms that Lévy envisions that the development of the network society¹¹ will lead to “the improvement of people’s lives and the fulfillment of their potential” (Lévy, 2011, p.143). Benkler supports this position when he explains that the current situation has reshaped how and who is responsible for cultural production. For him, the network society has enabled “individuals and groups to participate in the production of the cultural tools and frameworks of human understanding and discourse” and through this new culture how we interact with one another (Benkler, 2006, p.275).

Such potential has often taken the form, as Lévy indicated, in the exchange of production roles. Castells states that the network society has provided the ability to transform the binary pairs of author/reader and user/producer into easily shifting positions (2001, p.26). This fact is supported by van Loon: “There is no longer a single process of mediation”, between reader and author and text “but instead a continuous process of remediation” (2000, p.309). From this perspective, the relationship has been transformed. In fact, users and producers form spectrums of one and the other. The importance of this transformation of roles is summarized by Askehave & Nielsen when they explain that “the multi-medianess of web texts supplies the texts with a rich polysemous potential where the web user is ‘invited’ to participate actively in assigning meaning in the process of text consumption” and text production (2005, p.125). In other words the user of the Web is both a reader and a writer and has been afforded the ability to make, share, and interpret meaning.

¹¹Lévy’s main project is to provide a scientific representation of the cognition of the collective intelligence that is captured by the data streams that flow through the network society. The point of such a mirror is to provide a space where we are able to reflect upon and improve the cognitive processes of our collective intelligence. He actualizes this representation by creating an abstract machine (more correctly, an Information Economy MetaLanguage [IEML]) that automatically calculates the relationships between meanings. Such a map will allow us, “to identify unexpected similarities, complementarities that cut across categories and systemic gaps—which a labelling system limited to the usual classifications of disciplines and occupations would not have brought out” (Lévy, 2011, p.13).

2.3 Power & Conflict

As these two previous subsections have shown, the revolutionary changes in the material conditions of information and communication has affected “the entire realm of human activity” (Castells, 2004, p.9). Furthermore, these changes can be witnessed in the new relationships of production, power, and experience that permeate throughout our social lives (Castells, 2000, p.376). As a result, we are presented with a new situation, and with such a situation comes new dynamics in power. Castells explains that power “is the most fundamental process in society, since society is defined around values and institutions, and what is valued and institutionalized is defined by power relationships” (2009, p.10). He defines power as “the relational capacity that enables a social actor to influence asymmetrically the decisions of other social actor(s) in ways that favor the empowered actor’s will, interests, and values” (2009, p.10). Because power is exercised through the coercion of “will, interests, and values”, it does so by constructing meaning “on the basis of the discourses through which social actors guide their action” (Castells, 2009, p.10). Unlike previous notions, Castells explains “that the power-holders are networks themselves. [...] they are humans organized around their projects and interests. But they are not single actors [...] since the exercise of power in the network society requires a complex set of joint action that goes beyond alliances to become a new form of subject, akin to [Latour’s] action-network actor” (2004, p.32). It is through this avenue, of understanding how power reconfigures the network society, that offers insights into how the contours of society are aligned along some lines and not others.

But of course, power is no simple matter. As Castells explains the network society “is not absent of contradictions, social conflicts, and challenges from alternative forms of social organization” (2000, p.387). This happens because “each network defines its own power system depending on its programmed goals” (Castells, 2004, p.31), and thus “social conflicts take the shape of network-based struggles to reprogram opposite networks” (Castells, 2000, p.695). He explains that reprogramming comes in the form of the “scripting” of new codes or values that reconfigure the performance and goals of the network (Castells, 2000, p.695). This tactic of changing the codes is particularly effective because “once the rules are set, they become compelling for all nodes in the network, as respect for these rules is what makes the network’s

existence as a communicative structure possible” (Castells, 2009, p.43). Such reprogramming ultimately leads to the transformation of cultural structures, practices, and codes (Castells, 2000, p.697). Power thus manifests itself in cultural transformations. This fact is important because culture “is a collective construction that transcends individual preferences” and influences the “the practices of people in the culture” (Castells, 2001, p.37). In other words, what is at stake in network conflicts is the ability to contribute to one’s own culture.

As Castells explains, these conflicts happen at the level of discourse as they “frame the options of what networks can or cannot do” (2009, p.53). Because power is thus carried by discourse, power “in the network society is communication power” (Castells, 2009, p.53). Therefore the shape of discourse, and subsequently power, is determined by “one specific technology: communication networks that organize socialized communication. Because the public mind [...] is ultimately what influences individual and collective behavior, programming the communication networks is the decisive source of cultural materials that feed the programmed goals of any other network” (Castells, 2009, p.53). It is thus through discourse that the goals and the values of a network are actualized. Hence, conflicts between different value systems in the network society are effectively conflicts between networks, and their confrontations are made present within the space of discourse.

In order to investigate how knowledge has been transformed by the values the network society, the current thesis will investigate a network conflict within a particularly active project that incorporates many of the characteristics of the network society. For Tapscott, Myspace, Youtube, Linux, and Wikipedia are “today’s exemplars of mass collaboration” (Tapscott, 2006, p.11). Of these, Wikipedia has been a particularly common example (Lévy, 2009, p.91; Shirky, 2010, p.98; Benkler, 2006, p.70). Benkler goes so far as to say that “Wikipedia is the strongest example of a discourse-centric model of cooperation based on social norms” (2006, p.104). Considering the argument Castells has made that communication power is exercised through discourse, Wikipedia is an excellent candidate to explore how meaning production and communication power, through technology, has transformed the production of knowledge.

3 Wikipedia

In 2001, Jimmy Wales launched Wikipedia, an online encyclopaedia that describes itself as the “encyclopedia that anyone can edit” (Wikipedia, 2002). Within the last ten years it has surpassed its competitors in terms of popularity as it is currently the sixth most visited website on the Internet (Alexa.com, 2012a). Contrastingly, Britannica.com which is its closest encyclopedic competitor, is ranked a distant 6,185th (Alexa.com, 2012b). Because of this popularity and its user-generated content, Wikipedia has been positioned as a model project of the Internet age. Among its virtues, cyberculture scholars have described it as open, transparent (Benkler, 2006, p.289; Lessig, 2004, p.70), and an exemplar of mass collaboration (Tapscott, 2006, p.11; Shirky, 2008, p.109). Lévy goes so far as to describe it as “the most complete and most consulted encyclopaedia in the world” (2011, p.279). All of these sentiments are evident in the statement by Benkler that Wikipedia has created a “radically new form of encyclopaedia writing” (Benkler, 2006, p.70). This description of Wikipedia encompasses the perception that Internet projects are successfully creating novel forms of social organization, practices, and products that are important agents of societal change.

3.1 Wikipedia as wiki

Much of what has allowed Wikipedia to reach its acclaimed status has been due to its server infrastructure, also known as “Wiki”. This software was first developed in 1995 by Ward Cunningham, who defines it as “a freely expandable collection of interlinked Web ‘pages’, a *hypertext system* for storing and modifying information — a *database*, where each page is easily editable by any user” (Leuf & Cunningham, 2001, p.14, original emphasis). As wikis are based on hypertextual documents, wikis benefit from their flexible characteristics.¹² These features of wikis allow for a great degree of transparency to the production of content as it captures every single edit that has occurred in the history of a page.¹³ This allows for anyone to view

¹²For Tapscott, the web of links and the software that supports it disrupts the professional and credentialed authorities that once monopolized the production of both knowledge and culture. Lévy saw the potential for hypertexts in the early 1990s “as a textual machine that could profoundly change writing, and therefore thought” (2011, p.8).

¹³Returning to van Loon’s description that networks disrupt chronology and Castells’ timeless time, Wikipedians interact somewhere between synchronous and asynchronous time. In some cases the conversation happens in rapid succession, like in the case of edit wars. Viéga, Wattenberg & Kushal (2004) aptly describe how edit wars between Wikipedians plays out over time. Most other times the interactions can refer to any recorded comment at any time in

records on who has contributed to an article and what changes have been made. Ultimately these recording and archival features facilitate the generation of talk about a specific page. Each wiki page has its own “talk” or “discussion” forum where people propose changes and new directions for pages, argue the merits and value of certain information, and try to negotiate conflicts when they happen.¹⁴ Given that every conversation and edit is recorded through the wiki software, Wikipedia has become a space where emergent coordination can be observed as it happens.

An additional benefit of the wiki software is that it blurs the line between author and reader, a characteristic that is attuned to the abilities of the network society. For example, Ruth & Houghton state that the sharing “of authority is central to a wiki epistemology, as is empowering participants. Any user can participate in the creation of shared documents, which evolve through shared community goals” (Ruth & Houghton, 2009, p.137). By providing these activities to users, wikis “facilitate the construction of knowledge [...] such that interacting with learning processes occurs in place as a construction rather than a body of knowledge to be absorbed” (Ruth & Houghton, 2009, p.135). With these aspects in mind the authors outline that these characteristics result in a set of values that are the epistemological paradigm of wikis:

1. Collaboration – individuals acting together to develop shared knowledge
2. Construction/co-construction – individuals acting together to produce knowledge and their products (in flux)
3. Different ways of learning – individuals acting together as equals – sometimes an expert, sometimes a novice, rather than in competition
4. The authority of ‘the’ expert is undermined

the archives and project it forward as a new comment to contemplate and investigate. As Biuk-Aghai & Lei (2010) mention, the majority of communications between contributors is asynchronous. But what makes this system work well is that each communication, whether it be a comment, an argument, or merely a single edit, is archived.

¹⁴An area of research that has yet to be followed is the actual level of transparency that Wikipedia affords its users. While all articles record all changes and editors, it is almost entirely inaccessible to the causal reader. One must learn a new language and navigate through a data stream that is clumsily chronologically organized. Some articles do have topical organization of the arguments but for larger articles (for instance, the article Canada has 18 Archives of discussions, each contain between 20000–30000 words), there is a great limitation in terms of navigability. The lack of being able to dynamically sort through arguments and consensus limits the elasticity that the site has otherwise created. Additionally, not all conversations about articles are held on Discussion pages. As Pentzold & Seidenglanz (2006) outline, there are alternative places for discussion such as the Internet Relay Chat (IRC) channels that are dedicated to Wikipedian issues. This further calls into question the transparency of Wikipedia.

5. A different philosophical underpinning which is more oriented towards constructionism

(Ruth & Houghton, 2009, p.148).

In addition to these values, a sixth can be added. Ward Cunningham states that “the original Wiki was to have a point of view and promote that point of view” which he reiterates is a positive force of activism (Youtube, 2012). Cunningham is correct in this assessment as the values of collaboration and co-construction are centered around practices of self-identification.

To elaborate on these values, the following subsections review a number of studies of Wikipedia that have been conducted in order to shed light on how these epistemic values have been actualized on Wikipedia.

Self-identification: The network society, and by association, wikis, is made real by the individual. Benkler explains that individuals are now “better able to do things for and by themselves” (2006, p.130) and what they do is collaborate. Lévy describes that the “process of collaborative production of shared memory favours individual learning insofar as the individuals involve their personal experience in the conversations [...] and involve the results of the conversations in the reorganization of their personal experiences” (2011, p.101). In keeping with this position, Ruth & Houghton describe that wikis are devices that help individuals form knowledge by way of process rather than an already storehouse of knowledge (2009, p.148). Because it is a process made between individuals, Lévy states that it leads to “mutual recognition and enrichment of individuals” (1994, p.13). On Wikipedia this process is mediated through an article’s talk page as well as personal user pages and user talk pages “where it is customary to post personal messages” (Bryant, Forte & Bruckman, 2005, p.3). These are subsequently the main spaces for Wikipedians to “establish an identity within the community” (Bryant, Forte & Bruckman, 2005, p.7). Additionally, even though it is difficult to receive credit for one’s contributions on Wikipedia, users retain a sense of ownership over their work (Bryant, Forte & Bruckman, 2005, p.4) and are sometimes recognized by their peers for their efforts.¹⁵ While Wikipedia has a rule that all project pages have a neutral point of view (NPOV), this practice has been considered by

¹⁵On Wikipedia barnstars are awards given to Wikipedians by their peers to recognize hard work. They are displayed prominently on a user’s page.

some users to mean “the possibility that all perspectives can be accommodated in a community” (Westerman, 2009, p.139). This perspective is in keeping with the mutual respect that can only be achieved through individual points of view and self-identification.

Collaboration: Collaborative efforts are negotiated between Wikipedians through a number of self-conscious social norms, open discourse, trust in good faith, and agreements to meet some form of consensus (Benkler, 2006, p.72). Benkler explains that such collaboration of millions of users is made possible by modularity, which “maximizes their autonomy and flexibility to define the nature, extent, and timing of their participation in the project” (Benkler, 2006, p.100).

Wu, Harrigan & Cunningham explain that when a large number of Wikipedians collaborate on an article, they tend to “produce excellent authoritative articles” (2011, p.1). Likewise, these researchers observe that poor articles are found to have few contributors. This has also been found to be true by Wilkinson & Huberman (2008) who explain that the quality of an article positively correlates with the number of contributing editors. Additionally, they also find that the articles of the highest quality are often those that are the most popular. These studies are therefore describing how the increases in individual collaborations has resulted in successful steps toward realizing the goal of making an encyclopaedia.

Co-construction: Numerous studies have examined how users construct/co-construct Wikipedia through the self-organizing activities they participate in, such as consensus building. Beschastnikh, *et al.* (2008) report that consensus appears to be structured in many cases by the policies and rules that Wikipedians have already agreed upon. Likewise, Viégas, Wattenberg, Kriss & van Ham (2007) find that the construction and cohesion of the community is emphasized through discussions about group coordination, policy, and process. Laniado, Tasso, Volkovich & Kaltenbrunner (2011) supports this position when they study the different forms of discussions that take place. The researchers find that networks of direct replies consist of different conversational activities that outline coordination efforts depending on whether users are interacting directly with inexperienced users or mentoring and correcting behaviours to keep discussions appropriate to Wikipedia. Other users retain close-knit conversations that deal with specific article creation and maintenance issues. The authors explain that such behaviours stem

from article discussion pages requiring more social organization and role-assignment, sparking specific modes of communications between users. Likewise, Clark, Ruthven & O'Brian Holt (2009) find that what is communally important is reflected "in the implicit structures [of] the objects they create and share" (2009, p.5). Given this, the researchers discover that instead of a stabilization of genre, they see greater activity dedicated to the evolution of the encyclopaedic genre. These changes indicate shifting attitudes about what purpose the encyclopaedia plays in Wikipedians' lives. These studies show that it is through the act of discussion that users co-construct Wikipedia.

Cooperation: Picking up on Ruth & Houghton's description that learning in the wiki environment is through cooperation, Bryant, Forte & Bruckman (2005) uses the concept of a community of practice to explain how this occurs. A community of practice (Wenger, 1999) defines a group of people who collaborate around a single activity that carries with it a history of the way things are done. In much the same way that Laniado, Tasso, Volkovich & Kaltenbrunner (2011) describes mentoring activities through discussion, Bryant, Forte & Bruckman (2005) use communities of practice to describe the process of transforming a reader into an editor of Wikipedia. They find that the relationship between experienced Wikipedians and new users is one founded on the explication of rules and policies that these users can refer to and augment, creating the ground rules for cooperation on Wikipedia.

Trust in the community: While Ruth & Houghton describe the epistemology of wikis as the undermining of the "authority of 'the' expert", the current thesis translates it as a relocation of trust, from the expert individual to the active community, hence the renaming of this epistemic value. In this regard, Bryant, Forte & Bruckman explain that the first wiki-based encyclopaedia made by Wales originally "relied on editing and review procedures that were developed in print, expert-produced, for-profit publishing systems. The Wikipedia founders' brilliance was in their ability to identify the shortcomings of the traditional model and adapt their project to new constraints" (Bryant, Forte & Bruckman, 2005, p.5). One of the ways to circumnavigate the shortcomings was to use a wiki that disrupts "status hierarchies" by replacing the value of expertise with the value of activity (Ruth & Houghton, 2009, p.145). In this way "prior

training or education do not convey a priority status on the contributions of any one individual” (Myers, 2006, p.8). In place of a hierarchy of status, projects like Wikipedia “are based on a hierarchy of meritocratic respect” (Benkler, 2006, p.105). Likewise, Reagle relates that the key to the understanding leadership in an “open content community” such as Wikipedia is to understand that leadership is validated through merit (2007, p.144). Further to the point, activity is more important than expertise because it allows for “more open, potentially fluid interactions between participants” (Ruth & Houghton, 2009, p.137). This fact is supported by the fact that while anyone can become an administrator, the role can only be obtained through election. As such “one must have accrued the requisite credibility in the community to receive unanimously positive votes” (Forte & Bruckman, 2005, p.4).

It follows then that the authority of expert contributors has been subverted by the fact that “all contributors (even if they are anonymous) are held to be equal in epistemic authority” (Sanger, 2009, p.64). In this way wikis “are by nature anti-authoritarian, socially constructed, role shifting spaces that are used to create ‘shared knowledge’” (Ruth & Houghton, 2009, p.144). Therefore, truth claims “are addressed by the force of the argument, not by the credentials of the individual” (Myers, 2006, p.8).

Many early critiques of Wikipedia recognize this fact and follow-up with the question of how the veracity of Wikipedia can be ascertained if expert contributors are not given privilege. Fallis counters this position by stating that we “typically trust a particular encyclopedia entry not because we trust its author but because we trust the process by which the entries in the encyclopedia are produced. And the process by which entries in Wikipedia are produced seems to be fairly reliable” (2011, p.304). Therefore trust in knowledge claims has moved from the expert individual to the active community.

Constructionism: Seymour Papert defines constructionism as a view of learning that is established by “building knowledge structures” and that this happens “felicitously in a context where the learner is consciously engaged in constructing a public entity” (1991, p.1). This is certainly true of how Wikipedians use policies. As Butler, Joyce & Pike explain, the meaning and identity of Wikipedia is “reflected in discussions of policy changes that center on whether or not something is consistent with the ‘core principles’ of Wikipedia” (2008, p.1105). As

noted by Beschastnikh, *et al.* (2008) and Viégas, Wattenberg & McKeon (2007) the practice of policy citation is an act of socialization, a type of action that solidifies the community. Westerman uses Dorothy Noye's definition of a group to explain a similar process. He states that a group "emerges in the dialogue between a 'community of the social imaginary that occasionally emerges in performance' (the group of registered Wikipedians themselves) and 'the empirical network of interactions in which culture is created' (the activity of not only writing an encyclopedia but formulating the policy and an epistemology of how that encyclopedia is to be created)" (Westerman, 2009, p.146). What this illustrates is that as Wikipedians come together to learn they continually thread together the identity of Wikipedia.

As is evident in these epistemic values, wikis like Wikipedia create a social space where individual opinions and cultural expressions are enveloped by a set of discursive practices. These practices are guided by epistemic values that are in keeping with the characteristics imbued by the network society. In this way Wikipedia embodies the flexibility (it can easily add and subtract new nodes), scalability (it can be used to connect both small and massive groups), and its survivability (its distributed structure makes it impossible to shut down).

3.2 Wikipedia as encyclopaedia

According to Wikipedia:Five pillars (2005a), it describes itself as an "online encyclopaedia" that acts as "a comprehensive written compendium that contains information on all branches of knowledge" (Wikipedia, 2007b). Similarly, Featherstone & Venn define an encyclopaedia as "a device which systematically organizes knowledge of a known, or knowable world" (2006, p.5). Additionally, Emigh & Herring (2005) rely on the *Merriam-Webster Online Dictionary* for their definition of encyclopaedias: "[a] work that contains information on all branches of knowledge or treats comprehensively a particular branch of knowledge[,] usually in articles arranged...by subject" (p.1). As Wikipedia defines its project as encyclopaedic in nature, its existence relies heavily on what an encyclopaedia sets out to achieve. For some scholars, this poses fundamental problems in terms of Wikipedia's ability to be a new process of sharing knowledge.

Featherstone & Venn point out, even if Wikipedia utilizes a relatively new medium, Wikipedia continues to follow "traditional disciplinary divisions" and is "low on inter-disciplinarity and innovation" (2006, p.10). Additionally, the scholars warn of old habits like the typical exclusion

of non-Western knowledge, which they accuse Wikipedia of practicing (Featherstone & Venn, 2006, p.2). They continue their cautionary tone when they warn that as knowledge and theory travels quickly, the “democratization of knowledge” still remains in the power of dominant centres and, “authorizing procedures that continue to favour a specific western perspective” (Featherstone & Venn, 2006, p.3).

In alignment with these statements is the study conducted by Emigh & Herring (2005) who complete a textual analysis study of three online encyclopedias to understand how similar or different expert-generated and user-generated encyclopaedias are. They compare articles from the Online Columbia, Everything2, and Wikipedia. Their results show that writing on the Columbia and Wikipedia “are not significantly different from one another. Statistically speaking, the language of the Wikipedia entries is as formal as that in the traditional print encyclopedia” (Emigh & Herring, 2005, p.5). They continue by describing that both the “Columbia Encyclopedia and Wikipedia entries are systematic, standardized, and narrow in scope” (Emigh & Herring, 2005, p.9). In their discussion section, the authors ask “How is it that the wide-open participation structure of a wiki can reproduce traditional print norms?” (Emigh & Herring, 2005, p.9). They conclude, that like Crowston & Williams (2000), online encyclopaedias reproduce older forms of the genre. The authors thus state that such developments are paradoxical as “users who faithfully appropriate such systems create homogeneous entries, at odds with the goal of open-access authoring environments to create diverse content” (Emigh & Herring, 2005, p.1). They continue to explain that experimental users and the content they create may be removed by more active and normative users. Thus “Wikipedia users appropriate norms and expectations about what an ‘encyclopedia’ should be” and continue genre reproduction and assume “control over the content produced within the system, literally erasing diversity, controversy, and inconsistency, and homogenizing contributors’ voices” (Emigh & Herring, 2005, p.9). Indeed, these scholars point to the curious fact that the radical changes in the network society, here represented by “wide-open participation” did not result in a new form.

3.3 Wikipedia as a site of conflict

If the question of this thesis is to understand how knowledge is transformed in the network society, the preceding scholarship poses an apparent problem for using Wikipedia to understand

this process, as both Featherstone & Venn and Emigh & Herring argue, Wikipedia has done no such thing. In fact, it reproduces traditional genres and as such, produces knowledge in the same fashion. However, given the fact that Wikipedia is a wiki (which includes its own set of epistemic values) the issue seems to be more complex than a mere reproduction of a genre. The existential question then is how can Wikipedia both be a model project of the network society yet exhibit traits that run completely counter to these revolutionary values? From the opposite position, how can Wikipedia reproduce a traditional genre while upsetting the very grounds that make an encyclopaedia what it is? This issue is substantially confounding and leaving this issue unresolved will lead scholarship about Wikipedia, and knowledge in the twenty-first century, toward an intellectual impasse.

However, all is not lost and the differences within the literature can be explained by returning to Castells' discussion of power and conflict. Reiterating what has been previously quoted, Castells explains the network society "is not absent of contradictions, social conflicts, and challenges from alternative forms of social organization" (2000, p.387). Indeed, the literature about Wikipedia makes it quite clear that there are some existential contradictions at work. He explains that such contradictions and conflicts occur because "each network defines its own power system depending on its programmed goals" (Castells, 2004, p.31), and thus "social conflicts take the shape of network-based struggles to reprogram opposite networks" (Castells, 2000, p.695). Conceivably this is what is happening on Wikipedia, that the contradictions between Wikipedia being a wiki and it being an encyclopaedia are the manifestation of a network conflict.

It is the hypothesis of the current thesis that Wikipedia is indeed a site of network conflict and that understanding this conflict will lead to a better understanding of knowledge is transformed in the network society. In order for this hypothesis to be followed, at least two networks need to be identified. Considering that the literature points to divisions being made between Wikipedia being a wiki and an encyclopaedism, these two participles of Wikipedia's namesake are to be considered to be the conflicting networks. Additionally, the hypothesis also requires that the values and goals of these two networks be established. However, this requirement poses a problem. While the values of wikis has been made explicit through their epistemology, the same cannot be said of encyclopaedism. As far as the researcher has been able to discern, there has not

been a reciprocal study that adequately defines the epistemological values of encyclopaedism. In other words, further research is needed to establish the epistemic values of encyclopaedism. As such, the purpose of this thesis is to reveal the epistemic values of encyclopaedism in order to better understand the conflicting values present in Wikipedia.

4 Summary of the Literature Review

So far the question of how do changes in society and technology affect the production of knowledge has followed a line of inquiry that started with a discussion of Heidegger, Latour, and Castells' conceptualization of technology. In this discussion it was concluded that technology is the process of making relationships with the world around us. In this way, the contour of the network of relationships determines all that we need to know and to act. Through this lens, technology mediates the conditions of our reality. As such, Castells describes that through the confluence of three independent processes, (changes to the economy for management flexibility; the globalization of capital, production, and trade; the primacy of individual freedom and open communication (Castells, 2001, p.2)) our society has never been more dependant on the structure of the network. In the network society, the logic of the network dictates that power is established through communication. Therefore an empowerment of a network's actors manifests itself in the ability to contribute to the construction of one's own culture.

One such site of empowerment is Wikipedia. It has facilitated the ability for the mass collaboration of millions of users to not only create an encyclopaedia but also create a space where cultural meaning can be argued, negotiated, and constructed. Because of the epistemic values of wikis as described by Ruth & Houghton (2009) (self-identification, collaboration, co-construction, trust in the community, and constructionism) Wikipedia has actualized on a grand scale many of the characteristics of the network society. However, Featherstone & Venn (2006) argue that instead of revolutionizing knowledge, Wikipedia has reinstated the power of traditional structures, effectively contradicting the radical nature of Wikipedia. The current thesis makes sense of these discrepancies by considering that these contradictions are in fact markers of a network conflict over epistemic values. In this case, the two networks are wiki and encyclopaedism. The goal of the current thesis is to study the foundations of this conflict to

create a representative sample of how knowledge has been transformed in the network society.

It is at this juncture that a problem stands in the way of effectively analyzing how Wikipedia manifests its conflicted epistemology. While the epistemic values of wikis have been explicitly made clear by Ruth & Houghton (2009) and a host of other researchers, the epistemology of encyclopaedism has not been similarly articulated. Therefore the current study's contribution to the field will be in defining what an encyclopaedia is and outline its epistemic values. In doing so, the epistemology of Wikipedia can be observed through the conflicts of between wiki and encyclopaedism.

Therefore, in order to achieve this goal, the current study returns to Castells' explanation that the conflicts between networks of values occur within discourse. As Castells describes, communication power expresses itself as discourse and within discourse there are metaprograms that ensure "that the recipients of the discourse internalize the categories through which they find meaning for their own actions in accordance with the programs of the networks" (2009, p.51). Thus the following section reviews the different kinds of methods to define and analyze discourse in order to provide a methodology suitable to reveal the epistemic values of encyclopaedism so as to compare them to wikis, establishing the basis of conflict between the two networks.

Part III

Methodology

5 Discourse

As the chapter on the network society reflected, the ability to construct meaning in the network society is determined by the power of the network to program its own goals and values, which are expressed through “communicative action” (Castells, 2009, p.12). Castells also states that “discourses of power provide substantive goals for the programs of the networks. Networks process the cultural materials that are constructed in the variegated discursive realm. These programs are geared toward the fulfillment of certain social interests and values” (2009, p.51). What those goals are, how they are achieved, and how they take form are all questions that can be answered through discourse analysis.

Castells defines discourse along Foucauldian lines, stating that discourses are understood “as combinations of knowledge and language” (2009, p.15). In Foucault’s own words: “Discursive practices are characterized by the demarcation of a field of objects, by the definition of a legitimate perspective for a subject of knowledge, by the setting of norms for elaborating concepts and theories” (1997, p.11). He continues by stating that discursive practices “take shape in technical ensembles, in institutions, in behavioural schemes, in types of transmission and dissemination, in pedagogical forms that both impose and maintain them” (Foucault, 1997, p.12). What this description reveals about discourse is that meaning is established by creating legitimate forms of communication and that such communications happen in many different places and in varying degrees. These facts have been recognized by researchers of discourse and have resulted in a variety of analytical methods.

5.1 Review of discourse analysis

The scholar Bhatia notes that “text analysis, conversational analysis, rhetorical analysis, functional analysis, and clause-relational analysis” have all been used by linguistic studies “to understand the structure and function of language use to communicate meaning” (1993, p.3). These kinds of studies are marked by their attempt to find such structures and functions of

discourse within the linguistic features of a text. Examples of these sorts of analyses are like Van Leeuwen's (1987) study to understand how journalism communicates news or like Mishler's (1984) study about what happens when doctors and patients talk. In each case, understanding discourse comes from understanding how the text is structured. Contrastingly, instead of looking within the text itself, some studies rely on analyzing discourse from outside of the text, observing how the context of production influences the construction of meaning. For instance, in the study by Byars & Meehan (2000), they examine how the meaning of "feminine" was constructed by the Lifetime cable television channel in the 1990s. Similarly, Levine (2007) used cultural studies to investigate "How the various aspects of the production process contribute to texts and shape possibilities for audience readings" of the soap opera *General Hospital* (Levine, 2007, p.135). These kinds of analysis set out to discover how audiences, economics, producers, and ideologies act in pivotal roles in the creation of a discourse.

Even though these studies vary in methodology, they each attempt to answer the question of discourse "Why do members of a specialist community write the way they do?" (Bhatia, 1993, p.1).¹⁶ As in Foucault's description of discourse and the variety of studies, the answer lies within a complex of discursive layers. Part of the issue then is that each discipline approaches discourse from a different angle and for different reasons. As such, Bhatia explains that within discourse analysis there are a number of metrics to which a study of discourse can meet.¹⁷ Of these differences he explains that they come from two main positions. The first analyzes discourse in terms of description, making explicit the particulars that go into creating one kind of discourse over another. Studies with this purpose tend to concentrate "on the linguistic aspects of text construction and interpretation" (Bhatia, 1993, p.1). The second strives to explain and "rationalize conventional aspects" of discourse, most often relying on context (Bhatia, 1993, p.1). Acknowledging these differences, Bhatia poses the idea that the concerns of researchers reflect four spaces of discourse: discourse as text, as genre, as professional practice, and as social practice (Bhatia, 2004, p.18).

¹⁶Although Bhatia uses the term "write", discourse for him refers in general "to *language use* in institutional, professional or more general social contexts. It includes both the written as well as the spoken forms" (2004, p.3)

¹⁷Bhatia describes five spectrums of discourse analysis: "Surface level thin analysis – Deeper/thicker analysis; Lexico-grammatical form – Cognitive move structures and rationale genres; Focus on textual patterns – Focus on socio-cognitive patterns; Linguistic perspective – Multidisciplinary, multidimensional, multi-perspective; Confined to textual space – Extend to tactical and social space" (2004, p.18).

5.2 Spaces of discourse

Traditionally aligned with linguistics, *discourse as text* “operates essentially within a textual space where the knowledge about language structure and its function, [...] is exploited to make sense of it” (Bhatia, 2004, p.20). Furthermore he states that this level of analysis concentrates on “properties associated with the construction of the textual product, rather than on the interpretation or use of such a product” and often ignores contributions to meaning-making by the audience (Bhatia, 2004, p.20). Further to the point, research that concentrates on *discourse as text* “often excludes any significant engagement with context” (Bhatia, 2004, p.20). The next space, *discourse as genre*, rectifies this situation by examining discourse both in terms of text as well as context. He states that genre accounts for “the way text is constructed, but also for the way it is often interpreted, used and exploited in specific institutional or more narrowly professional contexts to achieve specific disciplinary goals” (Bhatia, 2004, p.20). Thus buttressing *discourse as genre* is *discourse as professional practice*. Bhatia explains that genres “often operate in what might be viewed as [a] tactical space [for] established members of discourse communities” (2004, p.20). Because of this space, the knowledge of a genre by an experienced community members allows them to expertly maneuver within the genre in order to communicate personal goals that are not part of the goals of the genre. Discourse here is understood as how experienced members bend the rules, conventions, and goals of the genre to suit their own needs. Finally, *discourse as social practice* is the last space in the spectrum. It represents an interaction with discourse that takes “context much further in the direction of broader social context, where the focus shifts significantly from the textual output to the features of context” (Bhatia, 2004, p.20). In this space discourse is used to both mobilize changes or stabilize the “identities of the participants, the social structures or professional relationships” and “the advantages or disadvantages such genres are likely to bring to a particular set of readers” (Bhatia, 2004, p.20). For Bhatia, each of these spaces addresses the various perspectives evident in discourse analysis.

5.3 Approaches to discourse

Bhatia emphasizes that analysts using this framework of discourse must understand that the spaces are not mutually exclusive and that any study “will necessarily pay some attention” to each domain (2004, p.21). He offers three examples of how this framework can be operationalized. First the socio-cognitive perspective is of the socio-linguist who concentrates heavily on the social context and only a small portion on the textual spaces. The second perspective being pedagogical, is like the applied linguist who investigates the textual space, rooting out the rules of a discourse so that they can be taught to novices who are a part of a larger social context (Bhatia, 2004, p.21). The third perspective, generic, views discourse as “genre within a socio-cognitive space, and will pay some attention to textual features of language use, especially to textualization of some specific features of lexico-grammar, and textual organization, on the one hand, and to certain features of social practices, especially those related to professional practices, on the other” (Bhatia, 2004, p.21). These three perspectives (socio-cognitive, pedagogical, and generic) display the trends in discourse research that Bhatia has identified. Of these three, he argues that the generic perspective offers a beneficial middle ground for understanding discourse.

He advocates that this approach benefits from balancing how language is used in a social space as “it pays all the necessary attention to the relevant features of textual form” (Bhatia, 2004, p.21). This is an important observation as Giltrow remarks that privileging context, as the genre theorist Miller (1984) does, often leads to a disservice to the important role that form plays in the construction of a genre (Giltrow, 2009, p.15). Like Bhatia, she believes that analyses of genre would benefit from balancing understanding “the contexts and the function that genre plays in these contexts” with relating “how these contexts are actualized in form” (Giltrow, 2009, p.15). Kwasnik & Crowston also follow suit in this regard. They state that as a result of genres capturing both individual and community meanings of text “genre provides an excellent lens for discourse analysis [...]. This type of analysis strives to understand not only the words, per se, but the contexts in which those words acquire meaning” (2005, p.81). Bawarshi also agrees with these scholars that analyzing genre provides a valuable insight into discourse. He states that “genre is an ‘actualizer’ of discourse, transforming general discourse into a socially recognized and meaningful text [...]. It is genre, thus, that gives a text a social reality” (Bawarshi, 2000,

p.349). In other words, genre textualizes social practices as well as it socializes texts. Aligning himself with Halliday, Bawarshi goes on to explain that genres “do not just express or help communicants communicate [meanings] rather, genres mediate and maintain these meanings” (Bawarshi, 2000, p.352). Along with Bhatia, these scholars argue that studying discourse from the generic perspective allows the researcher to understand how meaning is constructed by equally addressing the textual and social aspects from which discourse is composed.

Discourse, for the purposes that Castells describes, is a spectrum of interwoven social practices and textual conventions. Bhatia describes that most studies of discourse constitute it in terms of four spaces (text, genre, professional, and social practice) and research it from three perspectives (socio-cognitive, pedagogical, and generic). Of these perspectives, Bhatia shows a preference for the generic perspective as it balances social practices with the textualization of those practices. As such, genre analysis proves to be a particularly beneficial perspective to understanding discourse and furthermore, how power is communicated through and transforms the network society. The following chapter will explain what genre is, how it works, and will identify some model studies that address the reciprocal relationships between text and social practices.

6 Genre

6.1 Theory

6.1.1 Convention

In a typical folk use of the concept of genre, one may use it to describe a category of novels or films or bands that comprise of a set of expected qualities. For instance, when the genres of “romance” or “western” or “punk” are mentioned, certain ideas are conjured up as to what sort of book, film, or album one will experience. The important aspect of this conceptual use of genre is that it latches on to a key feature that all genres have, they lead us to expect one thing and not another. Martin (2003) states that genres are a means of establishing “consistent patterns that make it possible for us to recognize and predict how each genre is likely to unfold, and so manage new information, and interact appropriately and strategically” (p.7). Santini states that genre is important “because it reduces the cognitive load by triggering expectations through a number

of conventions” and “genres can be seen as sets of conventions that transcend individual texts, and create frames of recognition governing document production, recognition and use” (2011, p.4). Erickson echoes this sentiment when he states that a “genre structures communication by creating shared expectations about the form and content of the interaction, thus easing the burden of production and interpretation” (2000, p.3). Repeating this sentiment, Bawarshi quotes Bazerman that genres “are frames for social action [...] They are locations within which meaning is constructed. Genres shape the thoughts we form and the communications by which we interact” (Bawarshi, 2000, p.358). The argument that each of these scholars is making is that genre regularizes the interpretation and use of a certain text. In other words, genres constrain the number of potential meanings that can be construed from a text. For Martin, Santini, Bawarshi, Bazerman, and Erickson, such constraints are efficient at quickly directing the audience toward specific meanings of the text.

The flip side of genre’s ability to effectively communicate meaning is that any meaning outside of the expectations of the genre are categorically ignored. For example, Fairclough laments how in view of “the immense diversity of events in the world” the genre of news reporting reduces these events to “the often rigid formats of news” (1995, p.86). For film critic Rick Altman, meaning within a genre is reduced to “a restriction of possible perceptions to a specific, limited set of relationships” (1987, p.3). The result of such relationships is that, as Altman argues, instead of the structure of genre helping individual texts produce meaning, genres are actually restrictive and should be regarded “as complex methods of reducing the field of play of individual texts” (1987, p.5). In this sense, viewers of a genre lose autonomy in their ability to interpret a given text, as such interpretative activity is already “usurped by the genre” (Altman, 1987, p.4). For Altman, the interpretation of a text has more to do with the ideological forces at work in the genre than the creativity of the creator (1987, p.5). In this way, genre becomes an efficient carrier of ideology.

While not imbued with the palpable vehemence toward ideology that Altman has, Martin incorporates ideology into his analysis of discourse by stating that “ideology is the system of coding orientations constituting a culture” (Martin, 1992, p.507), and that within a given culture, there are different ideologies that are maneuvering to redistribute power through the “semiotic evolution” of discourses and genres (Martin, 1992, 507). Likewise, Bhatia (2004)

reports that discourse analysts Sarangi & Slembrouck (1994) investigate “how discourse is used as a powerful instrument of social control [...], to establish identities, to communicate ideology, or to influence and maintain social processes, social structures and social relations” (Bhatia 2004, p.11). Elaborating on this, Berkenkotter & Huckin (1995) confirm that genre “conventions signal a discourse community’s norms, epistemology, ideology, and social ontology” (1995, p.4). Finally, Martin states that since “each text is produced interactively between speakers, and between writers and (potential) readers, we can use it to interpret the interaction it manifests. Since each interaction is an instance of the speakers’ culture, we can also use the text to interpret aspects of the culture it manifests” (Martin, 2003, p.1). In these scholar’s opinions, genre is used to signal a variety of discursive intentions, whether they be for communicating an ideology, for establishing social norms, epistemologies, or simply manifesting a culture.

In summation, the conventions used by a genre are deliberate acts to produce particular meanings. Such meanings are therefore representative of what a genre values, and since genres are part of a social milieu, these values are shared by the culture that produces such genres. Taken in this light “the process of communicative action” that Castells describes when he states that meaning “is constructed in society through the process of communicative action” (2009, p.12) could be considered to be the work of genre.

6.1.2 Evolution

Despite genre’s characteristics of convention and limitation it is also subject to change and adaptation. Genres are not just static containers ready to be filled with stable sets of meaning. As Miller’s (1984) rhetorical understanding of genre explains “genres change, evolve, and decay” (p.163). Kamberelis reiterates this when he states that although “genres are relatively stable cultural forms, they are also flexible. They evolve, expand, and get inflected in different ways over time. This occurs because the people and collectives who engage in material and discursive practices informed by genres do so within ongoing historical trajectories, thus blurring the distinctions between genres, creating sub-genres and new genres, and re-inscribing old genres in new practices” (1995, p.160). He supports this line of thought when he states that “neither old nor new community members ever learn genres once and for all; rather they must continually learn the generic ways of making meaning with texts that evolve within the ongoing social-

retorical activity of the communities” (Kamberelis, 1995, p.160). In other words, the evolution of a genre stems from the activity of text-making that “creates a dialogue between inherited potentials and current experience. Meaning is always reevaluated or re-accented in the processes of text production and reception” (Kamberelis, 1995, p.132). In a similar fashion, Lefevere quoting Glowinski describes that genres change: “the functioning of a genre cannot simply be reduced to the fact that norms or laws that should definitely be observed are given. These norms only function as absolutely necessary for a relatively short period of time. What is necessity in a certain historical period [...] can be mere possibility in another” (Lefevere, 1985, p.669). Following this train of thought, Martin explains that “all thriving cultures are continually elaborating their meaning potential by combining old variables in new ways and by making new distinctions to combine” (1992, p.506). In order to understand how this process occurs, he states that genre “makes it possible to monitor this dialectic of constraints and possibilities” (Martin, 1992, p.506). Therefore, as much as a genre owes its function to the conventions it employs, what those conventions are, how long they exist for, and how they are used are all susceptible to change.

6.1.3 Identity

Just as discourse theory and analysis has been approached from different disciplines, the same continues for genre. Genre theory has been used by a variety of disciplines (literary studies, critical theorists, cultural studies, functional and applied linguistics, sociology, communication studies, rhetoric studies, composition studies) to describe and explain what genres are, how they work and how we use them. Of all of these disciplines, the origins of genre stem from the literary analysis in Aristotle’s *Poetics* where he establishes the formal differences between narrative and dramatic poetry, such as the epic or the tragedy. Gérard Genette, a French literary theorist, explains that this traditional notion of genre is generally understood as a set of static criteria that one can apply to understanding how literary works differ. Genre in this sense is used to establish a system of hierarchically arranged genres and subgenres from the mutually exclusive “basic” genres (or modes) of lyric, drama, and epic (Genette, 1992, p.9). It is from this preoccupation with the classification and identifications of genre that genre theory has, until the last thirty years, been preoccupied.

Contemporary theories of genre have begun to assess how genres function within society as opposed to how they are categorized by some sort of a priori logic. Askehave & Swales (2001) note that until the 1980s, discourse analysis had emphasized “on categorization of discourses as members of particular classes” (2001, p.196), a practice that is a testament to Aristotle’s influence. They recount that “the great advances in recent years have not come from increased sophistication about the categorization of genres, and of subsequent more robust criteria for determining genre membership, but from a whole range of studies that have deepened and widened our understanding of the roles of discourse in contemporary society” (Askehave & Swales, 2001, p.196). However, while genre theory has moved away from a superstructure of categories, part of the goal of genre theory is to be able to define what exactly a genre does. As will be explained in the following paragraphs it is through the efforts to realize a “communicative purpose” that one possible system of identification of genres has been realized.

Askehave & Swales laud Carolyn Miller’s seminal article “Genre as social action” (1984) as a crucial turning point in the recontextualization of genre definition. She argues that genre can be defined as a typified social rhetorical action, or, as the recurrence of social actions governed by rules for connecting forms with substances.¹⁸ She states that her definition of genre is rhetorically sound because it is not centered “on the substance or the form of discourse but on the action it is used to accomplish” (Miller, 1984, p.151).¹⁹ Swales via Bhatia provides a similar definition of genre: “a recognizable communicative event characterized by a set of communicative purpose(s) identified and mutually understood by the members of the professional or academic community in which it regularly occurs. Most often it is highly structured and conventionalized with constraints on allowable contributions in terms of their intent, position, form and

¹⁸In outline form, Miller enumerates her definition: “1. Genre refers to a conventional category of discourse based in large-scale typification of rhetorical action; 2. As meaningful action, genre is interpretable by means of rules; 3. Genre is distinct from form: form is the more general term used at all levels of the hierarchy. Genre is a form at one particular level that is a fusion of lower-level forms and characteristic substance. 4. Genre serves as the substance of forms at higher levels; as recurrent patterns of language use, genres help constitute the substance of our cultural life. 5. A genre is a rhetorical means for mediating private intentions and social exigence; it motivates by connecting the private with the public, the singular with the recurrent” (Miller, 1984, p.163).

¹⁹Put into action, she mobilizes this theory with Shepherd when they analyze whether blogs constitute a new genre (Miller & Shepherd, 2009). She uses the recurrence of typified rhetorical actions to identify that in fact “the blog” is more medium than it is a genre. This is due to rhetorical actions of blogs that happen within particular situations that blogs afford, such as the space for a personal or political voice. In such a case, the personal blog and the political blog are defined as genres whereas blogging itself is the medium that affords the space to create such genres.

functional value. These constraints, however, are often exploited by the expert members of the discourse community to achieve private intentions within the framework of socially recognized purpose(s)” (Bhatia, 1993, p.13). Building on these definitions, Askehave & Nielsen (2005) explain that genres “are not only characterised by shared sets of communicative purposes” but also represent “the way to go about accomplishing particular communicative purposes” (p.122). This explanation makes an explicit link between the communicative purposes of a genre and the recurrent actions that are used to achieve those purposes.²⁰

In fact, as Askehave & Swales quote Johns (1997) “purpose interacts with features of text at every discourse level” (Askehave & Swales, 2001, p.195). Martin supports this claim when he describes how even within the differences of form and content between a single text and a range of texts, or between genre and discourse “meaning is constructed on all levels, backgrounding the form/content dualities” (1992, p.496). Returning to Bhatia, this position is a similar reflection of the four spaces of discourse, that regardless of which space an analysis concentrates on, the purpose permeates throughout each expression, whether it is social or textual.

Ultimately the goal of any genre analysis is to identify how the communicative purposes shape a genre and consequently the meaning that the genre interprets. To make such identifications, Bhatia describes that tactical choices made by writers who contribute to a genre make use of either discriminative or non-discriminative strategies (1993, p.19). Strategies in this sense are actions that “are generally used in order to make the writing more effective, keeping in mind any special reader requirements, considerations arising from a different use of medium or prerequisites or constraints imposed by organizational and other factors of this kind” (Bhatia, 1993, p.19). For the most part, he describes that these strategies tend to be non-discriminative, that is “they do not change the essential communicative purpose of the genre” (Bhatia, 1993, p.20).²¹ Bhatia explains that “[d]iscriminative strategies [...] tend to vary the nature of the genre

²⁰Askehave & Nielsen argue that “the major linguistic reflection of communicative purpose is in the staging structure by which a text of a particular genre unfolds” (2005, p.122). They describe that this sort of analysis helps “establish whether some expressions are more preferred, and therefore more genre-specific, than others” (Askehave & Nielsen, 2005, p.122). For these two authors, the communicative purpose of a genre is best identified in the staging structure of a text, because the staging structure is the textualization of recurrent rhetorical actions. Although this is a possible route to seeing the communicative purpose in action, it is not the only one, as Bazerman’s 1988 study of the scientific article genre shows.

²¹Bhatia provides the example that two texts from the same genre may differ due to the medium (a non-discriminative strategy of communication), as in print and television advertisements, but this difference does not affect the generic communicative purpose to advertise. Thus despite differences in medium, they both belong to the

significantly, often introducing new or additional considerations in the communicative purpose of the text” (1993, p.21). Because these strategies alter the communicative purpose, they are helpful in identifying genres and sub-genres. Because such strategies differ from the regular conventions, they do not go unnoticed by the discourse community. Bhatia explains that if this is the case and a tactical choice goes against the communicative purpose of a genre, then it “leads to opting out of the genre and is noticed by the specialist community as odd” (2004, p.25).

In this way we can see that the communicative purpose is directly related to the recurrence of the strategies that the community uses in the “construction and understanding” of the genre (Bhatia, 1993, p.21). It is for this reason that the conventions that are derived from the communicative purpose of the genre are the limits to what a genre cannot exceed. In contrast, any conventions that are used outside of the communicative purpose are subject to change. Describing genre in this fashion explains how a genre can retain integrity while it undergoes constant evolution.

In the first two subsections of genre, I have described that genres are processes for producing the conditions for interpreting specific meanings from a discourse. This process is indebted to the communicative purpose(s) that a genre establishes. The actions of the genre community to realize a communicative purpose are called discriminative strategies, and as such, are integral to the identity of the genre. Those actions that are not related to a communicative purpose are described as non-discriminative strategies, and their presence explains how a genre can continue to be the same genre even though it evolves. The following subsection will display these aspects of genre by examining two approaches to genre analysis that are exemplar of Bhatia’s “generic perspective”. In each case, they balance their study by analyzing how the social and the textual aspects of discourse are entwined within genre.

6.2 Examples of genre analysis

6.2.1 Bazerman and the evolution of the scientific article

Two scholars who can be argued to fit into Bhatia’s “generic perspective” of discourse analysis are Charles Bazerman and his impressive *Shaping Written Knowledge* (Bazerman, 1988) and Gérard Genette with his theories of textuality. In the first case, Bazerman states, similarly to same genre (Bhatia, 1993, p.20).

Bhatia, that discourse must include both texts and context: “Words and meanings dialectically define each other. The immediate implication is that one cannot understand language without looking at the contexts in which it is used to convey meanings” (1988, p.296). It is this perspective that guides Bazerman on his rather ambitious project of explaining how the genre of the scientific article emerged as a result of a perceived need to communicate scientific discoveries among scientists and how it evolved over time to adjust to new needs. In keeping with Bhatia’s generic perspective, Bazerman conducts extensive textual studies that are made comprehensible by thoughtful investigations on the social practices and communicative purposes of the genre. The breadth and depth of his study is evident in the number of different research methods he utilizes to reveal how this genre unfurled.²² In each section of his research he actively seeks out all levels of discourse, from writers like Sir Isaac Newton responding to his audience, to analyzing linguistic features, to interviews. In each case, Bazerman locates different textual and social spaces of discourse and examines how the communicative purpose is manifested in each. What his multi-faceted approach to genre analysis proves is the comment made by Johns (1997) that the purpose interacts with all levels of discourse.

As indicated by these studies, Bazerman aligns himself with Miller, when he defines genre as a social action “a sociopsychological category which we use to recognize and construct typified actions within typified situations” (1988, p.319). Furthermore, he clarifies that a genre “is a social construct that regularizes communication, interaction, and relations. Thus the formal features that are shared by the corpus of texts in a genre and by which we usually recognize a text’s inclusion in a genre, are the linguistic/symbolic solution to a problem in social interaction”

²²To start he compares the purposes and strategies of three different scientific articles (one on DNA, another of social scientists, a third on poetry). This study is followed by a descriptive analysis of 100 articles from 1665 to 1800 in the *Philosophical Transactions of the Royal Society* in order to examine the emergence of the experimental report in order to “identify the kinds of problems the genre was attempting to solve and how it went about solving them” (Bazerman, 1988, p.63). Next he examines how the temporal relationship between scientific books and relatively quick scientific journals affected the rhetorical choices in Sir Isaac Newton’s *Opticks*. Subsequently, he examines “how social situations structure communication events and forms of communication restructure society” in analyzing how the printing press played a crucial role in creating new social roles for scientists (Bazerman, 1988, p.128). Moving on to the form of the article he analyses how “a web of common theory” created embedded relationships between spectroscopic articles of *Physical Review* from 1893 to 1980 (Bazerman, 1988, p.157). He does so through a textual analysis of the number of citations, the article length, sentence length, and syntax, as well as the names of headings and the amount of text dedicated to those headings. The results of this analysis conclude that a strong theory about spectroscopic studies played a crucial role in structuring these textual features of articles which he calls “theory as a textually integrating force” (Bazerman, 1988, p.157). In addition to these previous studies, he also conducts ethnographic studies with several physicists to understand how they read and use scientific articles.

(Bazerman, 1988, p.62). While not explicitly using the term “communicative purpose”, he is describing its qualities. He observes that such solutions may ultimately be unattainable but still direct the genre (Bazerman, 1988, p.16). Study of how goals are followed reveal not only why certain choices are made but also what is accomplished and what is not (Bazerman, 1988, pp.13–15). In view of this, Bazerman later explains that the value of genre analysis is to recognize the presence of epistemological choices, to unveil them, and reflect on the consequences of such choices (Bazerman, 1988, p.258). In this view, the goals of a genre are evident in the choices (discriminative strategies) that the discourse community repeatedly relies upon. To summarize, Bazerman argues that all social actions have an intended goal. Describing this goal is the first step in understanding genre. In the case of the scientific article, Bazerman defines its goal as the creation of a corpus of knowledge statements.

Justifiably, a genre is not only defined by its goals but also through the recurrence of specific actions. Bazerman explains that activities are repeated “because individuals perceive situations as similar and make similar choices” (Bazerman, 1982, p.316). He paraphrases Miller when he summarizes that “genre is a socially recognized, repeated strategy for achieving similar goals in situations socially perceived as being similar” (Bazerman, 1982, p.62). While decisions are made by individuals, these decisions become institutionalized and codified by the “collective wisdom” of the genre makers (Bazerman, 1988, p.316). Bazerman takes this position in describing how the genre of the scientific article emerged. Given that the generic purpose was “to produce statements of knowledge” (Bazerman, 1988, p.22), scientists sought the need for a consistent form of argumentation within their articles.

Bazerman also describes that there is a reciprocal feedback between texts and how the texts are “used in situations, and even the character of the situations themselves” (1988, p.22). Bazerman explains that the first place to look for such feedback is within “the textual form of written knowledge” (1988, p.18). While the text itself is a primary source, Bazerman explains that the choices made in the text interact with the lived-in world. He explains that “the word draws on and ties together writers, readers, prior texts, and experienced reality to constitute the domain [of] symbolic knowledge” (Bazerman, 1988, p.18). For example “[w]riting and reading may take place in privacy and composure, and they may carry out distant social actions, but they are still highly contextualized social actions, speaking very directly to social context and social

goals” (Bazerman, 1988, p.22).

Bazerman expands these ideas into an operable methodology that consists of four contexts; subject matter (or lexicon), connected literature, intended audience, and the author (1988, p.24). Lexicon describes the kind of information that the text conveys. Connected literature, or citation and reference, is used as an indicator of relationships between texts. Observing the intended audience reveals the “attitudes the text assumes that the readers will have, in the types of persuasion attempted, in the structuring of the argument” (Bazerman, 1988, p.25). Finally, observing the author’s claims at work reveals how an “individual statement-maker [comes] to terms with reality from a distinct perspective” (Bazerman, 1988, p.26). These four contexts constitute the metric that Bazerman used for analyzing the scientific article genre. He mobilizes this framework when he examines the typified actions of twentieth-century physics articles.

In his analysis he describes how the intended audience consisted of scientists and this determined the kind of language and argumentation that was used (Bazerman, 1988, p.155). He relies on a number of quantitative metrics to indicate how the genre evolved. He measures for the average length of articles as an indicator for dramatic changes in the production of articles (Bazerman, 1988, p.162) and the number of references and citations to identify the connected literature. He also analyzes sentence length, syntax, and word choice to observe any significant changes to how authors wrote. Finally, he examines how the organization, argument, and epistemology was communicated in order to reveal the logic behind the writing. Bazerman suggests that the study of structural organization and argument reveals how “discourse is intimately linked not only to knowledge and theory, but to [...] beliefs about what can be known, how it can be known, in what form it can be expressed and how it should be argued” (1988, p.174).

To prove his point, he methodically outlines that such examination can be executed at three levels of data: “the formal division and section headings” “the proportion of space devoted to the various parts of the argument”, and how the text presents its argument and logic” (Bazerman, 1988, p.174). He explains that while his study aligns with the “traditional view that science is a rational, cumulative, corporate enterprise” it is “realized only through linguistic, rhetorical, and social choices, all with epistemological consequences” (Bazerman, 1988, p.183). For Bazerman the importance of genre analysis is to explain “why regularities emerge, evolve, and vanish

(1988, p.315). Bazerman's study of genre is exceptionally applicable to the current thesis as he not only outlines different locations where the communicative purpose shapes the discourse, he specifies that the value of genre analysis is to uncover epistemological choices. Given that the current thesis has set out to explore the epistemological network conflict within Wikipedia, Bazerman will provide an excellent source of methodology.

6.2.2 Genette's paratext as a threshold to interpretation

Bazerman's perspective of genre is quintessentially social constructionist in nature, understanding social practices by observing their textualized forms. Because Bazerman is engaged with discourse at the generic level (as described by Bhatia), he does so by positioning the importance of his study in explaining social practices and supporting his claims with analyses of text. Instead of working from social contexts down into the text like Bazerman, Gérard Genette (1997) does the opposite and concentrates on establishing genre based on the textual relationships one text has with another. However, what constitutes a text for Genette extends well into social practices.

As Allen (2000) reviews, Genette is steeped in structuralism as his theory of transtextuality attempts to create an ontology of the meaning-producing relationships of a text. However Genette's structuralism can be described as "open" as it "gives up on the idea of establishing a stable, ahistorical, irrefutable map or division of literary elements, but which instead studies the relationships (sometimes fluid, never unchanging) which link the text with the architextual network out of which it produces meaning" (Allen, 2000, p.100). This approach to textuality affords genre analysis something that Bazerman's approach does not: a more nuanced understanding of Bazerman's context of "connected literature". For instance, Bazerman regards connected literature only as the act of citation. Genette's theory opens up what it means to be a connected text and provides relational qualities to different kinds of connected texts. For instance, Genette describes citation as an action of co-presence or, intertextuality (Genette, 1997, p.1);²³ through the external expressions of the text, like author interviews or internal structures like titles and headings (paratextuality, 1997, p.3); through relationships of commentary (metatextuality, 1997.

²³As noted by Genette (1997, p.1), his conceptualization of "intertextuality" is significantly more constrained than the more popular understanding described by Julia Kristeva (1997, p.1). She describes formal intertextuality as the "transposition of one (or several) sign-system(s) into another" (Kristeva, 1986, p.111). Elaborating on this view, she concedes that there is the more common use of the term as "the study of sources" (the definition that Genette uses) which has nothing to do with the semiotic theory she uses.

p.4), the imposition of one text upon a previous one (hypertextuality, 1997, p.5);²⁴ and finally, the linking of one text to any number of others through a common discourse (architextuality, 1992, p.82). He defines the architext as the intersection of formal, thematic, and modal “determinations that are *relatively constant and transhistorical*” (Genette, 1992, p.78). This description of the constant and transhistorical nature of architextuality is similar to Bhatia’s description of the discriminative strategies as they are means by which a genre retains its identity. Genette’s main concern is developing a theory of textuality that can accommodate all of the influences that affect the reception and shape of a text, that are “outside” of the physical printed page. For instance, he states the “literary work consists, exhaustively or essentially, of a text, that is to say (a very minimal definition) in a more or less lengthy sequence of verbal utterances more or less containing meaning. But this text rarely appears in its naked state, without the reinforcement and accompaniment of a certain number of productions, themselves verbal or not,” (Genette & Maclean, 1991, p.261).²⁵ What Genette’s structural typology of texts affords is a precise attention to the different contexts that influence the meaning produced by a genre.

While context is important to understanding how a text is made, Genette argues that contexts find entrance into the text through the paratext. This aspect of Genette’s theory has been applauded because it opens up literary studies to address the outside forces that affect textual production and interpretation. Kinser explains that context “neither frames the text nor consists simply of verbal fragments that have been reassembled into a new text. How a text is interlaced with context is illuminated and to some extent defined by paratext” (1990, p.17). Likewise, Maclean (1991) advocates that the “signs and ‘fringes’ which accompany and surround the text itself are an important and frequently neglected aspect of textual production” (1991, p.273)

²⁴Again, like intertextuality, Genette’s hypertextuality is a constrained understanding of the more common cybernetic meaning. As Castells describes in *The Internet Galaxy* (2001), a hypertext is “an actual interactive system, digitally communicated and electronically operated in which all the bits and pieces of cultural expression, present, past and future, in all their manifestations, [...] coexist and [can] be recombined” (p.202). Genette defines hypertext as a unidirectional relationship that has little to do with the computational network of hyperlinks that make up the internet.

²⁵As witnessed here, Genette has a very traditional concept of text, in that it is a literary one, where image, film, or sound as text is out of his intention. Even more so, as McGann (1991) observes “such matters as ink, typeface, paper, and various other phenomena which are crucial to understanding textuality [...] fall outside [Genette’s] concerns because such textual features are not linguistic. But of course all texts, like all other things human, are embodied phenomena, and the body of the text is not exclusively linguistic” (1991, p.13). Paling relates that McGann argues that we should overcome “Genette’s excessive concern with the text itself” (2002, p.138). However, Genette’s theory could easily extend into understanding the network of relationships between film, sound, and image texts.

and concludes that Genette's paratextual approach reconciles these differences. Paling (2002) advocates that "Gérard Genette's idea of *paratextuality*, by including what is already richly incorporated into our methods for organizing information, and adding clear links to a vital body of humanistic theory, can help bridge the divide that often seems to separate the humanities from the study of classification" (p.134). The value that Paling sees in a paratextual analysis of text is due to Genette's description that these elements "provide *thresholds of interpretation* for the texts to which they are attached" (2002, p.134, emphasis original). Falling into the descriptions of genre previously discussed, Paling states that the "use of texts, in other words, always exists as the use of a contextualized artifact" (2002, p.138). For these reasons, Genette's textuality is firmly in keeping with Bhatia's generic perspective that draws together both texts and contexts.

While Genette does not outline an explicit definition similar to communicative purpose, he alludes to the fact that paratexts are implemented to achieve purpose. However, the downfall of Genette's theory is that he retains a modernist's ideal about the sanctity of the author. Despite his extensions into the periphery of textual production, Genette's "author", and occasionally the author's publisher, are the only actors that create paratextual elements. And so all communicative purpose is for the sake of the individual and does not involve a greater discursive community. However, Kinser who mobilizes Genette's textuality, argues that what "an author writes seems today less expression than communication. It is part of a conversation with companions, known and not known, contemporary and not so" (1990, p.183). Following Kinser's use of Genette, the current study takes liberties with Genette's authorial limitation, and argues that paratexts are not only applicable to the purposes set by the author and publisher, but are also thresholds of interpreting the communicative purposes of the larger discursive community.

As Maclean translates Genette's description of the paratext in *Seuils* (1987) it "constitutes, between the text and what lies outside it, a zone not just of transition, but of transaction; the privileged site of a pragmatics and of a strategy" (Maclean, 1991, p.278). Given the previous discussion about tactical choices made by authors as discriminative strategies, the paratext represents a site where such strategies are realized. As Genette notes via Paling, the role of "the production of a text is done 'precisely in order to *present* it, in the usual sense of this verb but also in the strongest sense: to *make present*, to ensure the text's presence in the world, its 'reception' and 'consumption'" (2002, p.141). These words signal that paratexts are part of a

larger social structure, beyond author and reader, that it is entangled in all the social relationships that make genre a powerful source of interpretation.²⁶

This final subsection on genre has examined two methodologies for analyzing the “generic perspective” of discourse. By first looking at the social contexts and then drilling down into the textual artefacts, Bazerman’s analysis of the evolution of the scientific article explains how the discourse of the scientific community addressed recurrent issues of how to communicate scientific knowledge. Over time, responses to this issue became conventionalized. As such, these conventions display the epistemological choices that the genre needed for its communicative purpose. Bazerman shows how genres socialize texts by always relating how different textual features and social practices intermingle with one another as they evolve to realize communicative purposes. Likewise, Genette shows that the meaning of a text cannot be found solely within itself and must be considered in conjunction with a variety of contexts. Genette’s approach to genre comes from the opposite end, placing the text as the primary starting point and the interpretation of the text stems from the constellation of contexts that inform a reader as to how to interpret it. The main difference between these two generic perspectives is that Genette does not fully embrace the social qualities of a text as much as Bazerman. For Genette, even though he describes that acts of commentary, production choices by the publisher and author, citation practices, and imitation practices are pivotal to interpreting a text, he regards such actions as textual objects, not social practices.²⁷ Bazerman is well aware of the textual relations between a text and its contexts. He states that at “the same time as the meaning of the text is seen to reside within the reader, that meaning is also seen to develop out of a web of relations with other texts” (Bazerman, 1988, p.236). Furthermore, the “reader reads not a single text, but an intertext which creates both the traces of language familiar and meaningful for the reader and the presuppositions on which the reading rests” (Bazerman, 1988, p.236). Even though their methods are different, this sentence stresses the similarities that Bazerman and Genette have

²⁶A perfect example of the paratext in action is in Hutcheon’s (1986) study of how when used by historians, the footnote “remains the central material mode of textually certifying fact” in the construction of histories (p.304). While Genette states that citation is a relationship used in terms of co-presence, Hutcheon argues that citation is also a threshold for interpretation. For him, citation not only makes another text present, but it also acts as a powerful tactic for making knowledge claims.

²⁷A caveat is that in Genette’s *Palimpsest* (1997), there is one line where there is an inkling that texts are social constructs: “I view the relationship between the text and its reader as one that is more socialized, more openly contractual, and pertaining to a conscious and organized pragmatics” (p.9). However, the inclusion of this single sentence is not enough to make the claim that Genette is a social constructionist.

in regard to the interpretation of a text. As previously mentioned, what Genette affords that Bazerman does not is a comprehensive understanding of the *quality* of relationships between the text and other texts. For example, when Bazerman studies headings, he does so merely as a reflection of the communicative purpose in action, whereas Genette describes headings, or “intertitles”, as “liminal devices and conventions” that acts as both frame and content (Genette, 1997, p.xiv). Of the different kinds of textual relationships, the paratext has been used in various studies to explain how liminal aspects of text act as thresholds to interpretation.

In order to understand how conflict between networks occur within the network society, the current study will analyze discourse through genre by mobilizing Bazerman’s social-to-textual methodology and partnering it with Genette’s textual-to-social theory of paratexts. The resulting analysis will therefore contain three sections, the first to determine the communicative purposes of encyclopaedism, the second to examine the social contexts as to reveal the epistemology of encyclopaedism, and third, to compare how encyclopaedism has traditionally been textually structured to how it is structured now within Wikipedia. This three-pronged approach is in keeping with Bhatia’s “generic perspective”, in that it strives to create a balance between the textual conventions and social contexts that are responsible for the emergence, maintenance, and evolution of discourse.

Returning to the question of this thesis, genre provides a valuable set of definitions to understand how discourse works, which in turns explains how a network reprograms the goal and values of opposing networks. In considering how communicative purpose guides a genre, genre may be the “metaprogram” that is needed to effectively program networks. As Castells states, a metaprogram “ensures that the recipients of the discourse internalize the categories through which they find meaning for their own actions in accordance with the programs of the networks” (Castells, 2009, p.51). Put another way, as Askehave & Swales comment, Martin famously explains that “genres are how things get done, when language is used to accomplish them” (2001, p.195). In this review, we have seen that genre represents not only the process of how a particular discourse is internalized within a community but also how it creates the conditions for discourse to be sustained.

6.3 Summary

Because Castells explains that network conflicts happen through discourse, the struggle over Wikipedia's epistemology is discursive in nature. As such, the values of the two conflicting networks can be seen at work in discourse. Given that the epistemic values of encyclopaedism have yet to be defined, the current thesis appeals for a discourse analysis to discover them. The literature review used Bhatia's descriptions of the four spaces of discourse (as text, as genre, as professional practices, and social practices) to provide an access point to studying discourse. Bhatia explains that when researchers approach these discursive spaces, they often do so from three different perspectives: pedagogical, socio-cognitive, and generic. Of these three, the generic perspective best represents discourse as a network of relations that stretch from the social to the textual. In this way, the question of how a discourse community communicates the way it does is attended to in a balanced fashion.

To further elaborate how genre theory conceptualizes discourse, the review explained that genres can be identified by their communicative purposes. The reason for this is that any convention or recurrent action that disrupts the communicative purpose opts out of the genre altogether. Accordingly, authors exercise tactical decisions that are either discriminative (aiding in realization of the purpose) or non-discriminative strategies that are subject to change. Analyzing discourse from this perspective not only accounts for the seemingly contradictory nature of generic conventions and evolutions, but it also sets the communicative purpose as central to understanding genre, which in turn, is central to understanding why a particular discourse is used to program a network. In other words, it is through genre that the purposes of a network are achieved.

Additionally, the literature review examined two particularly salient examinations of genre. Bazerman's investigation of the emergence of the scientific article illustrated that genre analysis is apt at revealing the epistemology employed by a discursive community. Such discovery is important because epistemological choices are evident in the tactics and strategies of achieving a communicative purpose. Put another way, epistemic values are the guiding principles of a discursive community and direct the flow of generic evolution. As well, Gérard Genette's theory of paratextuality provides a comprehensive understanding of how the meaning of a text is

derived through its textual relationships.

By combining these two approaches the current study will be able to adhere to the requirements of Bhatia's generic perspective of discourse. As well, such an analysis will be able to not only identify the genre of encyclopaedism and reveal its epistemic values. Having identified these generic qualities of the conflict, the current study will conduct an analysis of paratextual features to understand how these values are textualized within encyclopaedism. The resulting analysis will show the structural differences between a traditional encyclopaedia and an encyclopaedia interfered with a wiki. The outcome of this three-pronged analysis can be used to understand on what terms do networks conflict with one another and how such conflicts result in a combined epistemic value set, effectively transforming our social reality through technology.

7 Genre analysis

The main question of the current thesis is based on the question of how knowledge is transformed through conflicts in the network society. Castells explained that cultural meaning (here, meaning knowledge) is established in the network society through the communication power of networks to reprogram other networks to reflect its values and goals. As such, the foundation of conflict is between alternative value systems. In order to see how this occurs, the literature indicated that Wikipedia would be an excellent site of investigation. However, to investigate Wikipedia, the epistemic values of the conflicting networks need to be made explicit. In the case of wiki epistemology, this was addressed by Ruth & Houghton's (2009) article. In the case of encyclopaedism, there is a problem of defining what is exactly an encyclopaedia. Thus the purpose of the current analysis is to identify the genre and reveal its epistemic values. Following Castells' lead, the current thesis turned to discourse analysis and subsequently genre analysis to find an appropriate methodology. In reviewing these analytical frameworks, Bhatia explained that a well rounded analysis of discourse includes a discussion of the communicative purposes as they exist through both the social contexts and its textual instances. As such, the current thesis will conduct a generic analysis in three sections.

The first section sets out to define the communicative purposes of encyclopaedism. The second section will utilize Bazerman's four contexts to describe how these purposes have been

realized in the social contexts and conventions of encyclopaedism. Additionally, these contexts will reveal the epistemic values that guide the genre toward its communicative purposes. The third section utilizes Genette's paratext to compare the textual product of Wikipedia and a traditional encyclopaedia to validate how the epistemic values of encyclopaedism affected the structure of encyclopaedic texts. The conclusion of these three analyses will help describe how the conflicts between wiki and encyclopaedism has resulted in a unique configuration of epistemological values, which in turn describes how our social reality is changed by the technology we use.

7.1 Communicative purposes

The first step that is necessary for the genre analysis is to identify the communicative purposes of encyclopaedism. To do so, the analysis in Section 1 will be a literature review of scholars who identify what communicative purposes an encyclopaedia serves. The descriptions of communicative purposes will then be used as the criteria to judge whether a particular text belongs to encyclopaedism, which in turns validates the corpus used in Section 2 of the Analysis. In addition, Section 1 will briefly explain the conditions that gave rise to encyclopaedism in the first place, drawing from both antecedent genres and pertinent issues of knowledge management of the time.

7.2 Social contexts

As was discussed in the Literature Review, Bazerman explains that the value of genre analysis is to "understand why regularities emerge, evolve, and vanish" (1988, p.315). To do so, one can first locate these activities within "the textual form of written knowledge" (Bazerman, 1988, p.18). Thus using the results of Section 1 of the analysis, Section 2 will use the communicative purposes as the criteria for choosing the corpus of generic works to be examined.

As the literature will suggest, four encyclopaedias have been pivotal to the development of the genre: Chambers' *Cyclopedia*, Diderot & d'Alembert's *Encyclopédie*, the *Encyclopaedia Britannica*, and finally Wikipedia. It is these four works that will serve as the basis to understand how encyclopaedism has emerged from the social contexts that activate the communicative purposes. The framework used to analyze these four encyclopaedias follows Bazerman's metric

of four contexts; subject matter (or lexicon), connected literature, intended audience, and the author (1988, p.24). By using these contexts, the conclusion of this section of analysis will not only describe how the regularities of the genre “emerge, evolve, and vanish” but will also recognize the presence of epistemological choices, unveil them, and reflect on the consequences of such choices (Bazerman, 1988, p.258). In addition to Bazerman’s framework, the connected literature context will be discussed by utilizing Genette’s textual relationships to add a greater level of description. Of the relationships he describes, both intertextuality (citation) and metatextuality (commentary) will be used to reveal how one encyclopaedia links itself to the previous canon, or tries to be situated within the genre through comparison.

7.3 Textual structures

While Section 2 analyses how the social contexts affected the production of knowledge, Section 3 examines the textual spectrum of discourse to discover how the epistemological choices of the genre have structured the text itself. Frey (2000) explains that textual analysis is useful when there are vast amounts of data, for studying changes over time, for texts that already exist, and for data that appears in context (p.574). Additionally, Klaus Krippendorff (2004) explains that content analysis (which includes textual analysis) attempts to understand contexts in ways that allow for the processing of “texts that are significant, meaningful, informative, and even representational to others” (p.41). This process occurs through observing significant differences between texts. Krippendorff explains that frequencies of components indicates “what is”; the ordering of these components indicates “what is important”; and the proximity between components indicates, “what is related to what” (2004, p.52). This utility of textual analysis is displayed by Bazerman’s study of the scientific article, which we be used a model for the textual analysis portion of the current thesis.

He explains that “an examination of the structure of articles will reveal even more about the way discourse is intimately linked not only to knowledge and theory, but also to epistemology — beliefs about what can be known, how it can be known, in what form it can be expressed and how it should be argued” (Bazerman, 1988, p.174). Bazerman explores this form of generic-textual analysis when he investigates the structural outlines of spectroscopic articles of *Physical Review* from 1893 to 1980. To discover how the epistemology of these articles was structurally

expressed, he examines three levels of data “(1) the self-identification of the articles structure as embodied in formal divisions and section headings; (2) the proportion of space devoted to the various parts of the argument; (3) the texts themselves, to extract the mode of argument and the logic of presentation” (Bazerman, 1988, p.174). Interestingly, Bazerman’s first level of data explains that headings are not only useful textual divisions but are windows into the epistemological decisions of the discursive community. This description is similarly aligned with Genette’s theory of paratexts, that textual features are thresholds of interpretation.

Given that this framework can identify epistemological beliefs within the textual product of a genre, the current thesis will use it to compare the headings used in an encyclopaedia and compare it to those used in Wikipedia. This approach has been proposed to be a valuable avenue of research by Kafker & Loveland (2000) who state that tracing “the filiation of articles from edition to edition, as well as new borrowings from other sources, gives a sense of the process of composing an eighteenth-century encyclopaedia, not to mention the priorities of individual editors” (p.4). Similarly aligned, Crowston & Williams suggest that future research of encyclopaedias should analyze “the evolution of entries in online knowledge repositories over time” (2000, p.10), which they suspect will show “increasing formality and homogeneity across the lifespan of a Wikipedia entry” (2000, p.1). Therefore, given the descriptions by Frey and Krippendorff, and the suggestions of research direction by Kafker & Loveland and Crowston & Williams, a comparative textual analysis would provide the current research with a framework to explore encyclopaedic epistemology over time through frequency, order, and proximity. Therefore differences in headings will indicate where epistemological beliefs have changed, similarities will indicate the continuation of an epistemological belief system that continues to be rooted in encyclopaedism. The following is a methodological outline as to how the current study establishes an appropriate encyclopaedia to compare to Wikipedia and a representative set of articles from which to derive the data.

7.3.1 Procedure

Frey (2000) outlines the following procedures to conduct a quantitative content-analysis of a text (p.575–579): Selecting texts (Acquiring appropriate database of texts and acquiring a representative sample); Determining the unit of analysis; developing content categories; coding

units into categories; analyzing the data.

The following sampling of texts took place during the month of December 2010.

7.3.2 Selecting texts

As the Literature Review described, the values of the network society are exemplar in mass collaboration projects like Wikipedia. As such, the representative texts examined should in themselves be exemplars of the process of mass collaboration. Therefore, the appropriateness of each text will be judged by the amount of their collaborative activity. The current thesis chooses the language site with the most number of articles to represent mass collaborative activity on Wikipedia.²⁸ In this case, English Wikipedia is chosen because its 4 million articles far exceeds the next largest language, which is German with 1.5 million articles. Because of this fact, the traditional encyclopaedia that will be used for comparison must also be English. From the encyclopaedias studied in Section 2, the *Encyclopédie* cannot be used, which then leaves the *Cyclopaedia* and the *Encyclopaedia Britannica*. Of these two, the *Britannica* is the preferred choice for two reasons. The first reason is that it has existed for a longer period of time and thus encapsulates more encyclopaedic practices than the *Cyclopaedia*. Kafker & Loveland (2009) support this position when they state that if one is “exploring change over time in works of reference”, the *Britannica* is an excellent work to use for study (p.3). The second reason is that in addition to being an encyclopaedia from the print era, the *Britannica* has also been active on the Internet. As will be discussed in Section 2 of the analysis, the online Britannica has only taken advantage of the distributive affordances of the Internet but not some of the more radical advantages, such as the ones that Wikipedia embodies. As such it continues to retain the primacy of expert authority, a value that is not considered to be partial to the network society. Since the *Britannica* has operated in the media of print and the Internet without (arguably) adjusting its values, it is a perfect candidate to compare how specifically the wiki (and not the Internet in general) has interfered with encyclopedism on Wikipedia.

²⁸ While this approach offers an approximate understanding of the amount of collaborative activity, it does not directly measure collaborative activity itself. Perhaps a more accurate measure would be to calculate the size and number of interactions within each language’s associated talk pages. Of course, undertaking such a task would require time and research that are outside of the scope of the current thesis.

7.3.3 Acquiring a representative sample

There are several different ways of obtaining a representative sample. Neuman outlines the possible sampling techniques as probability sampling (simple random, stratified, cluster) and non-probability sampling (quota, purposive, snowball, and convenience) (2000, p.345-370). Of these, the most appropriate for the current study is a purposive sample. This kind of sample is beneficial for a number of reasons such as in exploratory research or in order to find unique cases (Neuman, 2000, p.347-348). Viably, the following is the procedure for establishing the final sample used for analysis, the goal being to create a database with ten representative articles.

Criteria for choosing appropriate Wikipedia articles for analysis: As stated before, the criteria for choosing Wikipedia articles is guided by their ability to represent Wikipedia as a project of mass collaboration. As such, an approximate way of determining such activity can be found by observing the number of revisions that an article has. Wilkinson & Huberman (2008) describe that high quality articles result from intense “cooperative behavior” and that such behaviour was accompanied by high numbers of edits and editors (p.23). While the number of edits is automatically updated for each page, the researcher was unable to find a similar automatic count of editors per article. As such, the researcher has relied on the metric of revisions to identify articles with high levels of collaborative activity.^{29 30}

While articles with the most revisions are representative of Wikipedia’s collaborative efforts, not all of these articles have a representative article in the First Edition of the *Encyclopaedia Britannica*. As Table 1 indicates, the top most edited pages tend to be specific to Wikipedia and certainly do not have comparable articles in the *Britannica*. Thus, one must look further down the list for articles that are present in both Wikipedia and the *Britannica*. Results: Table 2 (p.168) lists the 24 most edited Wikipedia articles that also exist in the First Edition of the *Britannica*.

Having identified which of the most revised Wikipedian articles have a comparable twin in the First Edition of the *Britannica*, the next step is to ensure that the Wikipedian articles were established in the first year of Wikipedia. This makes the study comparable in terms of the

²⁹Source: http://en.wikipedia.org/w/index.php?title=Wikipedia:Database_reports/Pages_with_the_most_revisions&oldid=405019708 December 31, 2010

³⁰Results: Appendix, p.165, Table 1, Edits.

lifespan of each encyclopaedia. Interestingly enough, all of the Wikipedia articles at this stage of the selection process all started within the first year.^{31 32}

Relying on the guiding principle of mass collaboration for selecting the texts, the next step is to identify the order of popularity of the remaining texts. This is an additional measure for identifying the amount of participation that takes place within an article. As Bryant, Forte & Bruckman (2005) explain, reading an article is peripheral participation that contributes to the “overall goal of the community” (p.2). They state that at “the periphery of Wikipedia, novice users contribute by reading articles out of interest [...]. For the novice, the goal of participating in Wikipedia is often information gathering (using the site as an encyclopaedia)” (Bryant, Forte & Bruckman, 2005, p.2). Similarly, Priedhorsky *et al.* use page views as an indicator of the value of an edit. They explain that if an edit is left unchanged for a substantial number of revisions, then it is viewed to have a higher value than those edits that are quickly reverted (Priedhorsky *et al.*, 2006, p.260–61). As such, peripheral participation through page views acts as passive consensus building. The statistics used for determining article page views unfortunately only exist for the years 2008, 2009, 2010. The results reflect this limitation.^{33 34}

Cross tabulation of popularity and editing activity: The actual numbers of views and numbers of edits per article have been recoded in terms of rank within the sample. This allows for the levels of popularity and editing activity to be combined for each article, resulting in a cross tabulation of the most edited and most read articles that are common between Wikipedia and the 1771 edition of the *Encyclopaedia Britannica*.³⁵

Taking stock of the current list: Of the 25 articles that are included in Table 2, 20 of them are countries. The other five are *London*, *Cat*, *Atheism*, *Democracy*, and *Human*. The issue that could arise from analyzing these articles is a redundancy in the kind of data that emerges from the texts. So in the service of establishing a corpus of representative samples, it is in the view of the researcher to find other common articles and to leave only two countries in the analysis. As such, 18 countries will be removed from the list at random.

³¹Source: <http://en.wikipedia.org/w/index.php?title=NAME-OF-ARTICLE\&dir=prev&action=history>

³²Results: Appendix, Table 2, p.168, DateCreated.

³³Source: <http://stats.grok.se>

³⁴Results: Appendix, p.168, Table 2, Traffic 2008-2010 and a sum total of those years (2008+09+10).

³⁵Results: Appendix, p.168, Table 2, Edits+Traffic.

- a) Through the act of random selection, the following country articles have been chosen for removal from the list: *India, Philippines, Iran, Japan, Germany, Russia, Scotland, Brazil, England, Greece, Turkey, Italy, France, Portugal, Spain, China, Sweden, Dominican Republic*.
- b) With this reduction, only two country articles will be used for analysis: *United States* and *Canada*.

Finding replacement candidate articles for the corpus: In order to cover a breadth of topics, three articles have been added, *Ocean, Black, and Cannon* to the corpus. They are the next most visited and revised articles that are not countries, were created in 2001, and of course exist in the First Edition of the *Encyclopaedia Britannica*. As Appendix, Table 3 (p.168) shows, these three articles received significantly less visits and revisions than the seven other articles.

Number of editions/versions of the articles for collection: There are fifteen editions of the *Britannica*³⁶ and the choice of studying how the structure of 10 articles change over 240 years is used to rule out temporal irregularities. For example, not all editions contain the exact article mentioned within the list, the name may change, or disappear from one edition but then reappear within the next. As such, not every edition is available to be analyzed by the researcher. Thus, Table 4 lists every edition that is possible to analyze for the purpose of this study.³⁷

Inconsistencies between Britannica's editions: Not every edition of the *Encyclopaedia Britannica* contains each of the ten sample articles. While on the whole, most articles are represented in each edition, some are left out from time to time, or appear within other articles. This inconsistency further validates the reason for including the full lifespan of the *Britannica*. Table 5 displays the presence, absence, or integration of the article within a particular edition of the *Britannica*. Table 6 displays the bibliographic information for each article chosen for the corpus.³⁸

Comparability between Wikipedia and the *Encyclopaedia Britannica*: *How to compare the temporal irregularity of editions of the Britannica with the temporal irregularity of revisions made to Wikipedia:* Given that editions of the *Britannica* are irregularly released and availability

³⁶Although there are fifteen editions, some of those editions have yearly updates, creating over eighty separate publications of the *Britannica*. See Table 4 (Appendix, p.169) for full publication information.

³⁷Result: Appendix, p.169, Table 5.

³⁸Result: Appendix, Tables 5 (p.169) and 6 (p.170).

further conflates the issue of accessibility, one must decide how to compare the temporality of the articles between encyclopaedias. One way of choosing the Wikipedia versions is by recoding the publication dates in terms of a ratio of the entire lifespan of the articles.³⁹ Put simply, an article from the *Britannica* that is published in 1910 is a ratio of 58/100 of the total history of the *Britannica*.⁴⁰ This ratio is then transposed onto the lifespan of a Wikipedia article, indicating that a specific version of a Wikipedia article should be chosen that is also 58/100 of its total lifespan. This metric creates a relative lifespan comparison between articles from the *Encyclopaedia Britannica* and those of Wikipedia. (For the sake of ease, the Wikipedia revisions, instead of using years to indicate editions, they have been assigned letters (A-N).⁴¹

If the Britannica article ratio indicates a specific day for a Wikipedia article, which revision is chosen if a) there are multiple revisions for that day, b) there are no revisions on that day, or c) the revision for the day was subjected to severe vandalism:

- a) If there are multiple revisions for that day, then the last daily edit is the chosen sample. The reason for this choice is that if there is an instance of vandalism and edit wars, a cursory examination shows that the last edit tends to be made by a person who reverts the vandalism.
- b) If there are no revisions for that day, then the next closest edit is chosen, regardless of whether it is after or before the optimal date.
- c) If there is evidence of severe vandalism (here defined as a byte change of over 50%) then a reverted version is used. Although, if there is not substantial vandalism, but has limited vandalism (under 50%), the article is still accepted into the sample.

³⁹The reason for making the two encyclopaedias comparable in terms of lifespan is to help the coding process. First, the *Britannica* sets the stage as only 14 editions are available to the researcher. Therefore it make sense that 14 editions would be chosen from Wikipedia. This being the case, the obvious issue is that Wikipedia does not exactly have editions the way that Britannica does. Instead of publishing an edition with all the changes made at once, Wikipedia uses a system of instant revisions, resulting in an articles that have thousands of “editions”. Therefore, in consideration of the *Britannica*’s 14 editions and Wikipedia’s thousands, which Wikipedia revisions should be chosen for comparison? One possibility could be to evenly divide the number of revisions by 14 and have the Wikipedia revisions equally distributed. However, the changes in the *Britannica*’s articles are not just differences between editions, but also differences in time. Some editions are more spaced out in terms of the lifespan of the *Britannica*. As such, to make up for this discrepancy, the Wikipedia revisions have been chosen to represent the same percentage of distribution to take into account the amount of time between the *Britannica*’s editions.

⁴⁰ $[(\text{Year}] - 1771)/2.4 \text{ [% 240 Years]} = \text{ratio number}$

⁴¹Results: Appendix, p.173, Table 7.

7.3.4 Determining the unit of analysis

Both Krippendorff (2004, p.103-106) and Frey (2000, p.576-577) recognize that there are typical ways of defining units for content analysis. They state that one can unitize for physical, syntactical, categorical, and propositional distinctions in texts. The purpose of these units is to operationalize theoretical concepts in order to offer a metric of comparability between texts. While these (physical, syntactical, etc.) are the usual forms of textual units, the literature review on genre has provided an additional unit type that is particularly salient for the current study: paratexts. Genette describes that a paratext surrounds and extends the text “precisely in order to present it, in the usual sense of this verb but also in the strongest sense: to make present, to ensure the text’s presence in the world, its ‘reception’ and consumption in the form (nowadays at least) of a book” (1997, p.1). He continues by stating that “the paratext is what enables a text to become a book and to be offered as such to its readers and, more generally, to the public. More than a boundary or sealed border, the paratext is, rather, a threshold” (Genette, 1997, p.1). In other words, the paratext functions as a mediating device that structures the creation and reception of a text.

Genette indicates that a paratext has a number of attributes which contribute to its ability to mediate a text: through time, place, substance, and pragmatic status (1991, p.269). The temporality of paratexts is an attribute that conceptually divides paratexts into two kinds; the epitext and the peritext. Genette describes that there are texts that occur prior to the release of an author’s text as well as afterwards. Examples of these are interviews, prefaces in subsequent editions, and press releases. These kinds of texts are called epitexts and exist in a timeframe that is external to the original text. The peritext on the other hand, exists in conjunction with the appearance of the original text. Examples include titles, intertitles, and copyright information.

In addition to time, Genette explains that a paratext also has a spatial relationship to a text. “An element of paratext, at least if it consists in a materialized message, necessarily has a *positioning*, which one can situate in relationship to that of the text itself: around the text, in the space of the same volume” (Genette, 1991, p.263). In addition, the attribute of substantial status is dependant on contextual factors that influence our reading of a particular work by a particular author, produced by a particular institution. Examples of this attribute exist in the knowledge of

the author's or publisher's age, sex, nationality, etc. This sort of context that exists outside of the text contributes to how the book is produced as well as how it is received by the audience.

The final attribute of a paratext is its pragmatic status and function. This function is characterized by "the nature of the sender and addressee, the sender's degree of authority and responsibility, [and] the illocutionary force of the sender's message" (Andersen, 2002, p.43). This function then is one that happens between the author, publisher, and the audience of the text. Of all the attributes, this function is the most important. As Genette describes "the paratext in all its forms is a discourse that is fundamentally heteronomous, auxiliary, and dedicated to the service of something other than itself that constitutes its *raison d'être*. This something is the text" (1997, p.12). In the end, analyzing texts through paratexts emphasizes how certain textual devices facilitate the reading and interpretation of that text.

The reason why a paratextual analysis is appropriate for the current study is thus: encyclopaedism has the fundamental goal of exposing knowledge through its system of categorization. This system permeates the entire encyclopaedic text; from its alphabetization, subject choice, and finally, how articles are internally organized. Such internal organization occurs at the peritextual level, especially through heads and subheads that are positioned within the text. This function aligns perfectly with Krippendorff's statement that the frequency, ordering, and proximity of textual units indicate "what is" "what is important" and "what is related to what" (2004, p.52), as well as Bazerman's argument that headings reveal epistemological choices. Indeed, paratexts like headings are the thresholds where encyclopaedic values are enacted to achieve the communicative purposes of encyclopaedism. That being said, collecting encyclopaedic headings for analysis will trace the epistemic values through textual categories.

The following outline describes how encyclopaedic heads qualify as paratexts. First and foremost, heads and subheads exist in the same timeframe as the original text, thus falling under the category of peritext as opposed to epitext.⁴² The substantial status of encyclopaedic heads relies primarily on the fact that heads in an encyclopaedia are used as a crucial device to indicate how knowledge is connected (as will be explained in Section 2 of the analysis). Depending on the encyclopaedia, the name, age, or nationality of the author of the heading may or may not

⁴² As Genette notes, epitexts, the partner of peritexts, occur outside of the text in a different timeframe such as interviews or private letters from the author (1997b, 5).

known. In the case of a known author, then the headings may be seen as part of the apparatus of an expert and may be viewed as reflecting the knowledge of that particular expert.⁴³ If it is an anonymous author, then the substance of the heading is deferred to an encyclopaedic voice, where the substantive status of the heading reflects that of the substantive status of the encyclopaedia itself.

In regards to space and position, heading hierarchy is partly established through positioning above, beside or within a text. The Canadian typographer Robert Bringhurst explains that in the case of a work that has “many layers or sections”, it may use several kinds of heads “to remind readers which intellectual neighbourhood they happen to be visiting” (2002, p.20).⁴⁴ The position of these heads often indicates importance. For example, heads may be centered, aligned flush with the left of the body, aligned flush with the right of the body, indented from the left side of the body, or even placed in the margin to denote different levels of rank.⁴⁵ *The Chicago Manual of Style* (1982) explains that subheads “except the lowest level, are each set on a line separate from the text, the levels differentiated by type and placement[...]. The lowest level is often set at the beginning of a paragraph, in italics and followed by a period, where it is referred to as a *run-in side head*” (p.27). As well, the 2010 *APA Style Guide* arranges head position as the most important starting with center, then flush left, and finally indented (2010, p.62). All of these heading ranks have a tradition of being designated by alphabetical level (Level-A, Level-B, Level-C, etc.), as in the *Chicago Manual of Style* (1982, p.27) and by Bringhurst (2002, p.65), or by number level (H1, H2, H3, etc.) in HTML (Berners-Lee, 1993, p.16). Hence, the position of a head is directly linked to its pragmatic function, to create a meaningful hierarchy between the different parts of an article.

In addition, heads serve the text in a variety of other ways. The audience can use heads to

⁴³This fact will be elaborated in Section 2 of the Analysis, but in some encyclopaedias, intellectual celebrities were hired to write articles. In such cases, the substantial status of the head is supposed to be known and act as a validation for the content.

⁴⁴Bringhurst also explains that there is no set number of headings that any text should have, only that one should use “as many levels of headings as you need; no more and no fewer” (2002, p.65).

⁴⁵“One way to make heads prominent without making them large is to set them entirely in the margin” which he calls a shoulder and not a head (Bringhurst, 2002, p.65). *The Chicago Manual of Style* (1906) explains that center-heads are “usually set in caps or small caps” and aligned horizontally in the center of the body (p.88). Side-heads are either indented from the left side of the body or run in, typically composed of italics, caps, small caps, or bold. A separate kind of head is called the cut-in head, which is a box of white space cut into the body, set two lines under the first two lines of a paragraph, and composed in a different face (*Chicago Manual of Style*, 1906, p.88-89).

help locate a specific topic within an article or to skim over the important aspects of a subject without reading the entire article. This function is described in Kools's *et al.* (2008) article which details how the hierarchical structure of a text is crucial "for readers to get an overview of the materials and find information easily" (p.834). This function highlights the reference role that an encyclopaedia plays. Due to this function, heads are essentially paraphrases of the content that they relate to and tend to be "kept short, succinct, and meaningful" (*Chicago Manual of Style*, 1982, p.26). Similarly, Williams & Spyridakis (1992) describe headings that "occur as short phrases or topical labels and announce superordinate content before the reader encounters the actual content. During encoding and processing, headings should prime a reader's prior knowledge and assist him or her in activating the appropriate hierarchical framework or text schema to accept new information" (p.64). Beyond being aids in searching and scanning, heads also outline what aspects of an article are deemed to be more important to understanding the topic. While this is communicated through position, it is also communicated through form.

While Genette is otherwise careful to discuss every other aspect and species of paratext, he overlooks the part that typography plays in structuring the reception of a text. The only mention that he offers is that "The typesetting – the choice of typeface and its arrangement on the page – is obviously the act that shapes a text into a book", and that "I will simply note that typographical choices may provide indirect commentary on the texts they affect" (Genette, 1991, p.2). The generalization that typography is only indirect in its affect on the text is simply incorrect. As Bringhurst remarks, the "typographer must analyze and reveal the inner order of the text" (2002, p.21). Such order is visually communicated through the variables of typography; stance, weight, case, position, alignment, scale, decoration, and colour.

In the case of heads, the form that they take play an important role in distinguishing one rank from another. Bringhurst explains that this hierarchy is the sociology of typefaces. For instance, the axiom that the "upper case has seniority but the lower case has the power" has held firm for over twelve centuries, representing the European "constitutional monarchy of the alphabet" (Bringhurst, 2002, p.53). Table 8 (see Appendix, p.175) is a reproduction of this monarchy as described by Bringhurst, giving precedence to those typographic inventions that have existed for the longest period of time.⁴⁶ In addition to the "sociology" of typefaces, the *Chicago Manual of*

⁴⁶As Bringhurst outlines: "Ornamental initials, small caps, and arabic figures were early additions to the roman

Style (1906), outlines that enumeration also creates a hierarchy of both lists and heads.^{47 48}

Therefore a head is distinguishable through its stance (roman or italic), weight (bold or normal), alignment (center, left, right), position (body, indented, margin), enumeration (1, 2, 3,...), case (capitals, small capitals, lower case) and punctuation. These characteristics as defined by the typographers and style manuals have assisted in collecting the data necessary for the current study.

Given these attributes, a heading (head, subhead, or intertitle) can be defined as a peritextual device that structures a text. It does so by exhibiting the superiority of one part of the text over another. Superiority of heads is determined by how essential a section of the text is to understanding the total article. This is represented through rank and order. Rank determines which heads belong to larger, more essential, parts of the text (as deemed by the author, or editor). This is visually distinguished through a typographic hierarchy which includes location, scale, face, case, punctuation, enumeration, and decoration to communicate rank. Order on the other hand is established through the linear sequence of the text. Those text sections that appear first are deemed more essential or important to understanding the article than those that appear last. These attributes just mentioned are used to carefully evaluate the presence and rank of any headings found within the collected data. For the sake of clarity, a number of liminal devices have been excluded from the study.⁴⁹

union” (2002, p.53). Italics existed as a separate entity until the sixteenth century when sloped caps, roman, italic, and small caps began to be designed as single families. Bold and condensed faces began to be first used during the nineteenth century (Bringhurst, 2002, p.54).

⁴⁷The sequence of enumeration is as follows: Uppercase Alphabet, Uppercase Roman numerals, Arabic numerals, lowercase alphabet, Arabic numerals, lowercase alphabet, Greek alphabet (*Chicago Manual of Style*, 1906, p.57).

⁴⁸In addition to the typographic forms mentioned, there is a rare but traditional form of emphasis called *sperrsatz*, or of adding additional spacing between letters. Initially, *sperrsatz* was used only with blackletter fonts, as they do not have italic or small caps companions to indicate emphasis. This practice became hotly discussed amongst nineteenth and twentieth century typographers who deemed the practice altogether “ugly” for use with lowercase letters. “The punch-cutter Unger, creator of Ungerfraktur (c. 1800) and a famous typographer, declared that Schwabacher was an ugly type and introduced letter-spacing for emphasis in fraktur (previously, Schwabacher had been used for emphasis in fraktur)” (Tschichold, 1928, p.77). As well, the typographer Frederic W. Goudy is often quoted as having said that anyone “who would letterspace lower case would steal sheep” (Erik Spiekermann & E. M. Ginger, 2003, p.7). It has continued to be used in combination with titles set in all capitals or small caps. Bringhurst advocates for this usage as it is claimed to increase legibility (2002, p.30–31).

⁴⁹*Author byline*: Does not communicate rank or order, only authority; *Figure titles and captions*: Do not paraphrase text but paraphrase or describe illustrations or photos; *Table titles and captions*: Do not paraphrase text but paraphrase or describe data; *Headings for article bibliography*: Do not structure the article itself, but structure the sources for the article; *Article index*: A device used for locating instances of keywords within a text which are not always a paraphrasing of a section of text; *Article table of contents*: A reflection of the order established by headings which makes it redundant for the current study.)

7.3.5 Developing content categories

Once the headings are identified, the following categories are recorded about each head.

Label: The words, numbers, punctuation used within the head.

Style: A code used to indicate the combination of typographic elements that are used to indicate the level of hierarchy. Stylesheets of these attributes are located in the Appendix, p.176, Table 9.

Rank: Directly related to style is rank. Style is a visual indicator of which headings are superior and inferior within a text. This function is thus called rank. Nonetheless, not every article is equal in its use of style. Some articles within an edition may use three heading styles, whereas another in the same edition may use five different ones. Rank is thusly determined within the article, so even though a style used may be H2, it may in fact, be the first heading used in an article, so it receives a rank of one. This allows comparability between articles and editions regardless of differing editorial decisions about style.

Order: The sequence that each heading appears in.

Position Percentage: The position ratio is used as a way to determine where within an article a head exists. For instance, does it occur at the beginning, in the middle or the end? To create this number, the total number of headings within an article are counted, divided into a hundred, and multiplied by the order number of the heading.⁵⁰ This produces a percentage that shows its position within the article.

Parent: The ID number associated with the heading that is one level superior than the current one. Recording this category will identify the hierarchical relationship between headings.

Location: The page a heading appears on.⁵¹

⁵⁰ $n(100/N)$ where n is the order number and N is the total headings in an article.

⁵¹This code was recorded to aid future researchers in finding a particular heading and not for direct analysis. In terms of when the location is an URL, researchers can copy the heading using the “Find” command in the browser.

Edition: The edition a heading appears in.

ID: The code value attached to each instance of a heading, recording which publication, article, edition, and order the heading belongs to.

All of these categories are established to analyze which words are the most frequent, the most important, and the hierarchical source of most other headings. The results of this analysis will provide two profiles, one for each encyclopaedia, that can quantitatively discern differences and similarities. Given that headings are integral to achieving the encyclopaedic goals of education and reference, the results of this study will aptly describe whether the organization of knowledge in these two encyclopaedias significantly differ, describing how Wikipedia is either different or similar to the epistemological beliefs of the Britannica.

7.3.6 Limitations

While a textual analysis will expose valuable information about how encyclopaedic writing manifests itself through headings, the current context of a Master Thesis does create certain limitations. For the most part, much of the data collected was transcribed by hand. Given that the current study has been conducted by only one researcher, there are doubtless typographical errors and unfortunate omissions within the initial transcription. To exacerbate this issue, some early editions are transcribed from poor reproductions. Future studies using the same methodology must employ several proofreaders to ensure the quality of the data and use higher-quality source materials. Additionally, as there has only been one researcher, there is no level of inter-coder validation to ensure that coding of the data was completed with the utmost of fidelity. If the results of this preliminary study show promise, future research will need to reevaluate the coding procedures and be conducted by a team of coders.

Part IV

Analysis

8 Section 1: Communicative purpose

8.1 Encyclopaedism

This study is not the first to consider encyclopaedias as their own genre. A number of historians like Fowler (1997, p.22), Ribémont (1997, p.47), Kafker & Loveland (2009, p.1), Yeo (2007, p.50), Doody (2009, p.2), and Jacob, Treves & Gage (1997b, p.63), all describe a genre called “encyclopaedism”. Doody advises that studying encyclopaedism through genre provides “a useful matrix for understanding texts partly because it downplays the question of historical contingency and gives readers a means of relating one text to another across time and place” (2009, p. 18). Jacob, Treves & Gage also insist that “writing the history of encyclopaedism and encyclopaedia implies historicizing the forms of the organization of knowledge, looking for differences and variations more than identity and permanences” (1996, p.44). Therefore, while observing similarities is important for establishing a generic commonality between these works, it is just as important to witness their differences. The theoretical positions of these historians are in keeping with Miller and Bazerman by emphasizing how encyclopaedias are the realization of generic goals within a particular social context.⁵²

8.1.1 The creation of a universal system of total knowledge

The starting point of the current study, as it was with Bazerman’s study of the scientific article, is to first identify the goals of encyclopaedism. A number of historians have proposed their own definitions that tackle such identification. For instance, Withers (1996) defines encyclopaedism as the idea and practice of ordering and classifying knowledge (p.276). He argues that the goal that is common to all encyclopaedias is “to present a comprehensive conspectus of the knowledge of their age” (Withers, 1996, p.278). In keeping with the social constructionist

⁵²As Bazerman describes, the alternative position to a socially constructed encyclopaedia is an encyclopaedia that is an aggregate of natural facts. Such an aggregate comes from the belief that “to write science is commonly thought not to write at all, just simply to record the natural facts” (Bazerman, 1988, p.14). In other words, in the last century especially, scientific language has been seen as “a transparent transmitter” of the natural world (Bazerman, 1988, p.14).

view, he reiterates that this is achieved through the situated social activity of classifying the world (Withers, 1996, p.278). The historian Loveland also describes encyclopaedism as the attempt at a complete and total classification system. He describes encyclopaedias as “works of reference aspiring to cover all knowledge” (Loveland, 2006, p.59). Additionally, total or complete knowledge is often discussed as the systematization of a unified knowledge. In this vein, Featherstone & Venn define the encyclopaedia as “a device which systematically organizes knowledge of a known, or knowable world” (2006, p.5). Historian Aude Doody reiterates this position by describing encyclopaedism as both a “philosophical impulse towards universal knowledge” as well as “a literary artefact of organized information” (2009, p.1). Given the comments of these scholars, the purpose of encyclopaedism is to expose the facts of knowledge through particular classification systems of all knowledge.

8.1.2 Educating the public through systems of knowledge

In addition to the goal of classifying the world, encyclopaedism also engages in transforming the knowledge of a few people into knowledge for the public. The reason that education and encyclopaedias are so closely linked can be explained by the etymological history of the words. The word ‘encyclopaedia’ stems from the ancient Greek words *enkuklios paidea*. This is the most popular place to start when describing encyclopaedias.⁵³ For Burke (2000, p.93), Yeo (1991, p.24), Jacob, Treves & Gage (1997a, p.3), and the encyclopaedist Coleridge (Sullivan, p.320) *enkyklios paedia* meant the “circle of learning”.⁵⁴ *Enkyklios paedia* is also often defined as a general, rounded, or liberal education (Katz, p.20, 1998; Doody, 2009, p.4; Stecchini, 1962 p.3; Yeo, 2007, p.48; Stover, 1962, p.36; Sullivan, 1990, p.317). Jackson states that it comprises the “basic subjects” or “the usual everyday education received by all” (1977, p.351). In general, these varying definitions of *enkyklios paedia* all emphasize some form of education.⁵⁵

⁵³Katz (1998), Hadot (1997), Jackson (1997), Lih (2009), Stecchini (1962), and Fowler (1997) each start their articles by describing the relationship between Classical Greece and its *enkuklios paidea* and the history of encyclopaedism in their writings.

⁵⁴However, historian A.P. Bos and philosopher Jeremy Bentham state that the use of the “circle”, denoting a “conclusive whole”, is a misleading concept. Bos argues that *enkyklios paedia* is composed of a series of domains that “form a hierarchic whole involving a step-by-step ascent rather than a circular composition” (1989, p.147). Jeremy Bentham also rejected the idea that the metaphor of the circle was appropriate to knowledge. While he offered an alternative metaphor, the field, he thought that any metaphor would, in the end, be inadequate (Yeo, 2007, p.56).

⁵⁵Ribémont mentions that an education through encyclopaedism is achieved through the transposition of “high-level and authorized knowledge to current and contemporary knowledge, with a particular orientation” (1997,

According to Doody, despite *enkyklios paideia* being attributed to the ancient Greeks, it is not until Quintilian's *Institutio Oratoria* (95 AD) that we find the words first recorded (2009, p.13). In keeping with the other definitions, he describes it to mean a general education (Doody, 2009, p.13).⁵⁶ The first century historian Pliny the Elder explains that his *Historia Naturalis* (77 AD) is the first book to contain all the subjects of *enkyklios paideia* (Katz, 1998, p.23), alluding to the general education being composed of a number of topics.

The first use of the single word 'encyclopaedia' is argued to be written somewhere between 1447 and 1514 when scribes placed the two Greek words, *enkyklios* and *paedia*, together (Fowler, 1997, p.27–29).⁵⁷ Following this usage, the first self-proclaimed encyclopaedia comes from Paul Scaliger's 1559 *Encyclopaedia, Seu Orbis Terrarum* (Doody, 2009, p.5). While there is an obvious connection between *enkyklios paideia* and encyclopaedism, there is a danger in drawing too close a comparison between the two. Doody mentions that since the writings of Henri Marrou (French historian, 1904–77), the recent trend for historians is to ignore such links (2009, p.4). Marrou argues that the word encyclopaedia "evokes a picture of universal knowledge, and however elastic it may have been, *enkyklios paedia*, never claimed to embrace the totality of human knowledge" (Doody, 2009, p.11). For him, the Greek words should only be considered as meaning a "general education" (Doody, 2009, p.11). This statement is affirmed by the 1842 Preface of the *Encyclopaedia Britannica* which states that while "the Ancients had the term, there were not amongst them any compilations resembling those to which the Moderns have applied it" (EB, 1842, p.vi). This position of separating the meaning of *enkyklios paideia* from 'encyclopaedia' is the one that the current study follows.

From these authors we see that the questions that encyclopaedism engages are epistemological in nature.⁵⁸ It assumes that knowledge can be classified, ordered, organized, managed, unified, systematized, produced, totalized, fixed in place, and taught. In summation, these actions are a response to the perception of an unordered and chaotic world. Moreover, the

p.53).

⁵⁶"I shall now add some concise observations on the other departments of study, in which I think that boys should be initiated before they are committed to the teacher of rhetoric, in order that circle of instruction which the Greeks call *enkyklios paideia* may be completed" (Quintilian, 78, 1.10.1)

⁵⁷Sullivan claims that the single word usage of *enkyklios paedia* occurred as late as 1559 (1990, p.317).

⁵⁸Epistemology or the theory of knowledge, asks "What is knowledge?" "Can we know?", and "How can we know?". The *Stanford encyclopaedia of Philosophy* expands epistemological questions to "[w]hat are the necessary and sufficient conditions of knowledge? What are its sources? What is its structure, and what are its limits?" (Steup, 2010).

communicative purposes of encyclopaedism are to establish a classification system that makes the chaotic world understandable and then uses that system as a medium for transforming esoteric knowledge into exoteric knowledge. Arguably, these goals alone do not define the genre in its entirety as they are shared by many other projects, like scholasticism and academia for example. While communicative purpose can be used to identify a genre, it is but a starting point as further clarification is needed to understand what is an encyclopaedia.

8.1.3 Generic membership

As indicated by Doody and Marrou, there is a tendency within historical scholarship to anachronistically claim that certain works and personalities belong to encyclopaedism. Doody's polemical article "Pliny's Natural History" (2009) convincingly argues that such claims attempt to establish "an overtly teleological narrative" (p.2). She aligns herself with historical analyses that reduce the weight of this narrative by tempering it with "caveats about the dangers of etymology and anachronism" as well as "a greater desire to distinguish between the philosophical impulse towards complete knowledge and the production of an encyclopaedic book" (Doody, 2009, p.2). In the case of the *Historia Naturalis* (77-79 CE), Doody explains that it was not written as an encyclopaedia "but encyclopaedism has been central to later understandings of the text" (2009, p.17). Agreeing with historian Richard Yeo,⁵⁹ Doody draws the line between encyclopaedias and antecedent works by beginning the timeline with eighteenth century self-aware encyclopaedias (Doody, 2009, p.4).

8.1.4 Antecedent influences

Doody is adamant about this distinction as she sees in her field the tendency to attach the label of encyclopaedism to texts that are not properly encyclopaedias. In many cases, the link being made is to genres and texts that were antecedently influential in the defining the emergent shape of encyclopaedism. As such, Jamieson (1975) states that antecedent genres "are capable of imposing powerful constraints" (p.414) and that in the creation of a new situation, old

⁵⁹"As Richard Yeo has put it, 'the encyclopaedia is ... closely linked with the emergence of modernity, with assumptions about the public character of information and the desirability of free intellectual and political exchange that became distinctive features of the European Enlightenment.'" (Doody, 2009 p.20)

genres⁶⁰ “may abide as a living presence” (p.406). As such, it has been noted that genres such as commonplace books,⁶¹ dictionaries,⁶² and summas⁶³ have played substantial roles in setting up what sort of text an encyclopaedia could be.

Beyond genres themselves, there are a number of texts that have influenced encyclopedias. Works like Pliny the Elder’s *Historia Naturalis*⁶⁴ (77-79 AD), Isidore’s *Etymologiae* (636 AD), as well as Vincent of Beauvais’s *The Great Mirror* (c.13th century AD) are often called forth as historical ancestors of encyclopaedias. In this same vein, a number of philosophers are sometimes mentioned as progenitors of encyclopaedism. For example, Plato (d.347 BC),⁶⁵

⁶⁰It must be noted that the genres that Jamieson is speaking of are specifically rhetorical genres. While study of rhetorical genres is a discipline unto itself, it still provides valuable insight into how genres work.

⁶¹One of the antecedent genres that has influenced encyclopedias has been the commonplace book. Chambers himself mentions that his encyclopaedia should be used as a kind of commonplace book, a genre of books that was used for recording and “selecting phrases, anecdotes, and ideas encountered in the course of reading” (Yeo, 2007, p.56). Burke notes that students during the sixteenth century were often advised to keep a commonplace book, a notebook which was “organized in systematic form, often in alphabetical order of ‘topics’ or ‘commonplaces’” (Burke 2000, p.181). As well, Burke describes that the practice of dividing a work up by the use of topics or heads stemmed from this tradition and is evident in the organization of the *Cyclopaedia* (2000, p.95).

⁶²In most instances, the dictionary is understood as an alphabetical organization of the words used by a particular discipline or language. As such there is no instructional order or connection between definitions with the exception of being alphabetically ordered. Kafker & Loveland describe that encyclopaedias differ from dictionaries because they involve a systematic organization of knowledge intended to educate readers (2009, p.80). As Burke recounts a similar distinction when d’Alembert states in the introduction of the *Encyclopédie* that encyclopedias can be either organized by the “encyclopedic practice’, in other words thematic organization, the traditional tree of knowledge” or the “dictionary principle” which is “alphabetical order of topics” (2000, p.184). However, even though there is this distinction the *Cyclopaedia*, *Encyclopédie* and the *Encyclopaedia Britannica* call themselves dictionaries in their subtitles and make substantial use of alphabetization to systematically organize their contents.

⁶³Summas are medieval manuscripts that provided readers with the foundation to study Christian scriptures. Durbin describes that these works were the “‘scientific’ systematizing of medieval law” in the practical form of teaching Christian discourse (1996, p.126). Kreeft defines the summa as “simply a summary. It is more like an encyclopedia than a textbook, and it is meant to be used more as a reference library than as a book. There is extreme economy in the use of words. There are no digressions and few illustrations. Everything is ‘bottom line’” (Kreeft, 1990, p.14). Kreeft explains that the “structural outline of [St Thomas Aquinas’] *Summa Theologica* is a mirror of the structural outline of reality” (1990, p.15). Additionally, Aquinas states that the “intent in this work is to teach everything that pertains to the Christian religion in a fashion suitable to the education of beginning students” (Durbin, 1996, p.126). In this way we see that the system that organizes the text is an educational tool in and of itself. Additionally, summas popularized the strategy of summarizing knowledge. However, unlike the early encyclopaedias, summas were constructed as summarized debate “because the Medievals believed [...] that dialectic could uncover truth” (Kreeft, 1990, p.14).

⁶⁴Sader cites that Pliny’s *Historia Naturalis* (77–79 AD) is often considered the first proper encyclopaedia (Durbin, 1996, p.125), however Doody argues that “the Natural History is an encyclopaedia precisely because people have read and used it as one” (2009, p.18) and not because Pliny wrote it as one.

⁶⁵Katz argues that Plato was the first encyclopaedist because he attempted to “synthesize all Greek knowledge up to his time or to create a type of encyclopaedia of various treatises that discussed what was then known about all areas of knowledge” (1998, p.42). Imbert also treats Plato’s work as an “encyclopaedia” (1997, p.11). There is actually quite a number of articles that argue for and against Plato’s membership to encyclopaedism. The argument stems from two competing definitions of *encyclopaedic*. There is one side that describes encyclopaedic knowledge as “knowledge of a multitude of scattered details” whereas “the acquisition of a general theory regarding the sciences founded on reasoning” was the goal of *enkyklios paedia* (Hadot *et al.*, 1997, p.38). The former intention has also been described as polymathic learning. As a number of authors indicate, positioning both Plato and the

Aristotle (d.322 BC),⁶⁶ Francis Bacon (d.1626 CE), and Gottfried Wilhelm Leibniz (d.1716 CE) are called encyclopaedists while they never created an encyclopaedia of their own. In addition to these influential works, there is a great variety of other works that have been described as encyclopaedias that enact the type of anachronistic dangers that Doody warns of.⁶⁷ Following Doody and Marrou, the study will only consider self-proclaimed encyclopaedias published since the eighteenth century. This temporal and nominal distinction is helpful in delimiting the scope of the current thesis. The following subsection will investigate the conditions that self-identifying encyclopaedias emerged from.

8.1.5 Emergence during the Eighteenth Century

As of the eighteenth century the technological changes of Gutenberg's printing press (1436) had changed reading and writing practices for over three hundred years. By this time, the publication

Greek concept of *enkyklios paedia* as encyclopaedic is a disputable claim. "The goal of the *enkyklios paedia*," according to Hadot "is not a knowledge of a multitude of scattered details, but [...] something of a comprehensive education" (Hadot *et al.*, 1997, p.39). Hadot *et al.* go on to say that polymathia "denoted a knowledge embracing the greatest possible number of bodies of knowledge" (p.39) and that "Platonism, [...] was not inclined toward polymathic or encyclopaedic studies" (Hadot *et al.*, 1997, p.48). Yet, Peeters and Stecchini agree that the term polymathia was indeed a term we would associate with encyclopaedic knowledge (Peeters, 2000, p.24). Peeters argues "polymathia neither creates happiness nor contributes to the achievement of intelligence" (2000, p.24). Alternatively, the ancients believed, as Hadot *et al.* indicate, polymathia was to be looked down upon as it described people whose main goal was "the accumulation of factual information" (Stecchini, 1962, p.4). While heated, the whole basis of these arguments are made moot by Aude Doody's thesis that a connection between philosophical figures like Plato and encyclopaedism should be categorically avoided.

⁶⁶Fowler comments that although Plato establishes the "unity of all knowledge" he does not do it with the "clarity, cogency, or thoroughness" of his student, Aristotle (1997, p.19). Peeters explains that the works of Aristotle are often understood to be "the first work of an encyclopaedic nature, as it seems to embrace and order all of the elements of knowledge" (2000, p.23). Durbin supports this claim when he says "an Aristotelian treatise embodies all the then-known views on a topic and incorporates them within a systematic treatment" (1996, p.125). He goes on to say that Aristotelian students collected their teacher's treatises with the intent "to produce a systematic summary of all the discussions on all the key intellectual topics of the day" (Durbin, 1996, p.125). Of all these claims for Aristotle's encyclopaedism, the most explicit one comes from the "General Preface" of the 1790 *Encyclopaedia Britannica* that claims with unabashed certainty that "the aim of Aristotle is to delineate the whole region of human knowledge, actual and possible; to point out the limits of every district; and to assign to every thing which can be an object of human thought its proper place in the vast arrangement" (EB, 1790b, v.1, p.ii). Aligning itself with Aristotle is surely a sign that the *Britannica* was engaging in the teleological narrative-building that Doody describes.

⁶⁷The following is a list of works that are sometimes described as encyclopaedias. The Egyptian, Amenemope, *A Text to Dispel Ignorance about Everything That Exists* (2000 BCE) (Jackson, 1977, p.342). The Chinese calligrapher Ouyang Xun's reference book *Leiju* (c.900 CE) (Chan, 2009, p.13). The scholar Ibn Qutayba's "encyclopaedia" of the tenth century CE (Featherstone & Venn, 2006, p.6). The Japanese "systematic handbook" created by the state to modernize and invent traditions. Indians during a similar era of modernization used "practical handbooks for modernity", that described current technological advancements (Featherstone & Venn, 2006, p.7). Nonetheless, given Doody's argument, it is still important to understand how these works may have been regarded by encyclopaedism as influences.

of books was an industry unto itself.⁶⁸ Considering that an encyclopaedia of the eighteenth century included tens of volumes, thousands of pages, cross-references, indexes, engravings, and multiple authors and editors, it needed the infrastructure of a book industry to support such an enterprise. While the book industry created the commercial environment for encyclopaedias to flourish, it was also responsible for the perceived need for encyclopaedias.

Within the first century after Gutenberg's press, there were already concerns about the inability to read all the available information (Yeo, 2003, p.61).⁶⁹ Many scholars of the period, like Gottfried Leibnitz who wrote in 1680 of that "‘horrible mass of books which keeps on growing’ so that eventually, he feared, ‘the disorder will become nearly insurmountable’" (Yeo, 2003, p.62). His contemporary Jan Amos Comenius was also so overwhelmed by the number of books that he wrote the *Patterne of Universall Knowledge* (1651), which proposed an "educational program" that dictated "that the unity of knowledge, once mastered by scholars, must be shared and hence simplified. This required an abbreviation of knowledge, some kind of ‘magnetic directory’ for people lost in a sea of books" (Yeo, 2003, p.62). This anxiety and response was also felt by Chambers, who claimed that "knowledge had to be reduced to essentials and collated in an accessible work" (Yeo, 2003, p.71). Encyclopaedias were then conceptualized as condensed libraries, which were in themselves, incapable of carrying all known books.

The book industry and information explosion were not the only factors that account for the emergence of eighteenth century encyclopaedias. Another factor was the development of new classification systems for knowledge. Withers (1996) describes that these works are modern artifacts where "classification was central to contemporary intellectual enterprise" (p.275). He begins by stating that the "idea was to use the accumulation of knowledge generated by many individuals working freely and creatively for the pursuit of human emancipation and the enrichment of daily life" (Withers, 1996, p.276). Indeed, both the *Cyclopaedia* and the *Encyclopédie* contained claims that through the unity of knowledge society would be "united by the power of science or reason" (Yeo, 1991, p.25). This particular understanding

⁶⁸Close to 1,000,000,000 books had been published by the end of the eighteenth century (Buringh, van Zanden & Jan Luiten, 2009, p.417).

⁶⁹The historian Burke's *Coping with Gutenberg* (2002) describes how Gutenberg had both released an explosion of books and a corresponding anxiety surrounding the overwhelming amount of information available in the early modern Europe.

of knowledge came from the classification scheme established by Francis Bacon's *Instauritio magna* (1620). Bacon's philosophy described that all knowledge could be divided into three categories; memory, reason, and imagination. Burke explains that if one looks at the seventeenth and eighteenth century curriculums, libraries, and encyclopaedias, they certainly suggest "that Bacon's reclassification was the most successful of the various attempts made at this time" (2000, p.99). Withers pushes the position further by stating that Bacon's work laid down "the philosophical principles of modern encyclopaedism" (1996, p.277). This claim is supported by the encyclopaedists Diderot and d'Alembert who adopted Bacon's philosophy for the structure of their encyclopaedia. Indeed, the "encyclopaedias, dictionaries and universal grammars were the rationalizing spirit of this age made real as text. These texts were legitimate social documents as well as surveys of intellectual enterprise" (Withers, 1996, p.280).⁷⁰ Each new encyclopaedia brought with it a wealth of knowledge that not only broadened traditional scholarly learning but reached out into other social spheres facilitated by its intentions of spreading rationality.

9 Section 2: Social contexts

9.1 Corpus of encyclopaedias

As the previous section indicated, the current thesis considers only self-aware encyclopaedias published since the eighteenth century as possible candidates for the generic analysis. In order to move forward in creating a corpus to examine the social contexts of the genre, the current study appeals to historians of encyclopaedias to delimit the list. While there are a number of encyclopaedias that get mentioned time to time, like Harris' *Lexicon Technicum* (1702–1710), Coëtlogon's *Universal History of Arts and Sciences* (1745), or the *Encyclopaedia Metropolitana* (1817–1845), the majority of scholarship concentrates on three main encyclopaedias. The first most referenced encyclopaedia is Chambers' *Cyclopaedia* (Durbin, 1996; Kafker & Loveland, 2009; Loveland, 2006; Withers, 1996; Yeo, 1991, 1996, 2003, 2007). Descriptions of the *Cyclopaedia* are usually dovetailed with its direct descendent, Diderot & d'Alembert's *Ency-*

⁷⁰By the mid-eighteenth century not only were there encyclopaedias, dictionaries, and universal grammars, but there were so many others kinds of encyclopaedic works that a dictionary of dictionaries was published. The *Table alphabetique des dictionnaires* (1758) lists among its 164 pages almanacs, glosses, etymologies, apostolic writings, thesauruses, lexicons, heraldries, genealogies, heresies, histories, vocabularies, canons, catalogues, tables, and indexes.

clopédie (Burke, 2000, 2012; Durbin, 1996; Kafker, 1973; Kafker & Loveland, 2009; Loveland, 2006; Withers, 1996; Yeo, 1991, 2007). The next work that tends to follow the *Encyclopédie* in terms of influence is the *Encyclopaedia Britannica* (Burke, 2000; Durbin, 1996; Kafker & Loveland, 2009; Loveland, 2006, Yeo, 1991, 2007). Finally, the recent appearance of Wikipedia has occurred in some of these same histories of the genre (Burke, 2012; Yeo, 2007). Based on the work of these historians, there appears to be an established canon of encyclopaedias that includes the *Cyclopedia*, the *Encyclopédie*, the *Encyclopaedia Britannica*, and finally Wikipedia as its newest member. Given this canon, the current thesis will use these four works to explore the social contexts and communicative purposes of encyclopaedism.

9.2 Chambers' *Cyclopedia*, 1728 CE

In 1728 Ephraim Chambers published his *Cyclopaedia, or, An Universal Dictionary of Arts and Sciences*. This two volume encyclopaedia was a compilation of scientific knowledge that Chambers single-handedly edited. Over the course of his life, he would work on five editions of the encyclopaedia, the final one being published a year after his death in 1740. In 1778, Abraham Rees added a supplement volume to Chambers' original two. Later, this supplement would be recreated as Rees own publication, the *New Cyclopaedia* (1820).

9.2.1 Lexicon

Chambers begins his encyclopaedia with a preface that claims that his system of cross-references is “perhaps the only way wherein the whole circle or body of knowledge can be deliver’d” (1728, p.i). This statement outlines Chambers' intended purpose to provide an access point to all knowledge through his textual device. This goal was motivated by the perception that the overwhelming abundance of books was too much for any one person. As Yeo remarks, “Chambers faced contemporary anxieties about the explosion of books and new scientific discoveries by contending that knowledge had to be reduced to essentials and collated in an accessible work” (2003, p.71). This claim is found in the preface where Chambers explains that his intention is to “answer all the Purposes of a Library, except Parade and Incumbrance,” (1728, p.ii). Chambers also claimed that his work was an archive of scientific knowledge, one that allowed the scientific community to review previous research, therefore avoiding unnecessary

repetition of inquiry (Yeo, 2003, p.71). However, given the space of only two volumes, this purpose would have to be executed by summarizing only the most essential knowledge of the most important subjects (Yeo, 2003, p.65). Chambers thus describes that “the difficulty lay in the form, and œconomy of it; so to dispose such a multitude of materials, as not to make a confused heap of incoherent Parts, but one consistent Whole” (1728, p.ii). These comments indicate that Chambers believed his book to be a complete summary of all the necessary terms for understanding a scientific world. The subjects of such a world are laid out in the full, and lengthy, title of the *Cyclopaedia*⁷¹ as well as the “Table of Arrangements” (See Appendix, p.163, Figure 1).

The contents of the table reveal a variety of subjects that the *Cyclopaedia* covered. At the highest levels Chambers divides knowledge into the scientific and the technical (or the natural and the artificial). Further on, the plan is divided into the sensible, rational, external and internal. The third level is divided among relations, quantities, powers, and structures. It is not until the fourth level of order that we begin to see the branches of science as individual topics. Given this plan, the first half of his order is philosophically derived. Yeo remarks that the inclusion of a plan of this sort was unusual for an alphabetically organized dictionary. “Chambers’s boast of going beyond previous dictionary makers in offering the chance of a ‘continued discourse’ was closely linked with this graphic display of the sciences” (Yeo, 2003, p.66). That being the case, the construction of this encyclopaedia is uniquely different than antecedent dictionaries.

However, like the *Lexicon Technicum* (1704–1710) that Chambers was inspired by, the *Cyclopaedia* is primarily organized by the alphabet. Chambers argues that topical organization of the sciences was both “artificial and imaginary” (1728, p.ii). As well, alphabetization provided the chance meeting between two ideas, where “numbers of things are thrown precariously together, we sometimes discover relations among them, which we should never have thought of looking for” (Yeo, 1991, p.28). Beyond alphabetization, Chambers created an innovative device, cross-references, to connect separated topics. Sullivan remarks that cross-references “points

⁷¹The full title of Chambers’ encyclopaedia: *Cyclopaedia: or, an Universal dictionary of arts and sciences; containing an explication of the terms, and an account of the things signified thereby, in the several arts, both liberal and mechanical and the several sciences, human and divine: The figures, kinds, properties, productions, preparations, and uses of things natural and artificial: The rise, progress, and state of things ecclesiastical, civil, military, and commercial: with the several systems, sects, opinions &c. among philosophers, divines, mathematicians, physicians, antiquaries, critics &c.*

a reader toward other entries in the encyclopaedia in the effort to create circles of discourse, circumscriptions of areas of related topics” (Sullivan, 1990, p.337). This scheme offered the reader connections between “terms” and various other articles, resulting in what Chambers calls “assemblages of ideas” (Yeo, 1991, p.41). Chambers argues that “multiple cross-references were essential to the constitution of encyclopaedic knowledge for, by connecting all the parts of the work, they transformed a compilation into a single treatise” (Sullivan, 1990, p.328). This approach enabled his intention of opening up the available parts of the work by providing “a course of references from generals to particulars” (Yeo, 1991, p.41). While Chambers allows the reader to navigate through the various links, he does so through categories which indicate the order that “they are most advantageously read in” (Loveland, 2006, p.61). Yeo declares that the articles chosen for publication “had the effect of privileging and consolidating the identity of scientific disciplines” (1991, p.39). This demonstrates that despite the objectivity that Chambers was trying to achieve through alphabetization, he inevitably designed a set of preferred readings. Regardless of such issues, Chambers’ invention of cross-references became an indispensable device in the organization of later encyclopaedias.

To summarize, the *Cyclopaedia* is a scientific dictionary in the sense that it limits its definitions to branches of science as shown by the “Table of Arrangements”. It is first organized alphabetically as a way of mediating the sheer amount of possible terms to be written about. Second, it is organized through cross-references which are tied together by a master plan that hierarchically organizes the book through a philosophy of knowledge, fulfilling the goal of a continued discourse.

9.2.2 Connected literature

Both Miller and Genette mention that genres often enfold other genres into themselves. The *Cyclopaedia* is no exception. While Chambers’ *Cyclopaedia* uses citations to reference where some of his information originates from, there is a large amount of text that is not sourced. This practice of copying without reference was common during this period, especially in the genre called common-place books. At the time, such copying was legal under the 1710 Statute of Queen Ann if the publication was dated twenty-eight years prior if the author was alive (Yeo, 2003, p.70). As well, Chambers reasoned that since his work served a public function, he

was allowed to copy verbatim from books that were even under copyright (Yeo, 1991, p.70). In his article “Plagiarism”, he remarks that no writer starts from nothing, that books are often the “Assemblages of other Peoples; and what they take from others they do it avowedly, and in the open Sun”. He finishes by stating that their “Occupation is not pillaging, but collecting Contributions” (Yeo, 2001, p.216). As such, the practice of copying verbatim without reference is a transtextual⁷² connection between the *Cyclopaedia* and common-place books.

In addition to common-place books, another transtextual connection is the dictionary. Chambers often describes his work as a dictionary and that other lexicographers had not attempted to provide a “continued discourse” within their works (1728, p.i). The most notable lexicographer that Chambers alludes to is John Harris (d.1719). The historian Collision recounts how Chambers had stopped his apprenticeship in map-making because he was so inspired by John Harris’ *Lexicon Technicum* that he had to create his own (1966, pp. 103-04). In turn, the *Cyclopaedia* can be understood in terms of a dictionary, like Harris’ book, one that similarly uses alphabetical organization to list as many technical terms as possible.

9.2.3 Intended audiences

In terms of readership, Chambers and the future editors of the *Cyclopaedia* came to the conclusion that there were two kinds of readers; educated readers and autodidacts. Yeo states that the educated reader would use the encyclopaedia to either explore unfamiliar subjects or as a reference from previous reading (2007, p.54). The autodidacts, or novice readers, were told by Chambers to turn to the “general name of the Science he wishes to explore”, otherwise they would be “cast out to sea without a compass because he did not know the parameters of various subjects” (Yeo, 2007, p.54). Yeo also reveals that Chambers encouraged the practice of adding more information or making notes in the encyclopaedia (2007, p.54). Kennedy includes the usage of cross-references as a way for each reader to write “their experience of an encyclopaedic volume as a nonlinear text” (2009, p.236). In this case, it shows that it was expected that the reader would interact directly with the text, forming individual paths, and making marginal changes to the work itself to make it even more useful.⁷³

⁷²Genette’s hypertextuality.

⁷³In brief, marginalia are the notes made in the margins of a book. The activity of creating such notes constitutes a readerly practice used to assist in understanding a text or remembering its contents. Glosses, rubrics, indexes,

Of the educated readers, specialist scholars and “those who are proficient in science” are mentioned by Chambers as intended readers (Yeo, 2003, p.68). These types of readers were envisioned to use the *Cyclopaedia* “as they would refer to a common-place book, in order to assist their memories, without the labour and the loss of time, which it would require to recur to a great number of distinct treatises, whence their knowledge was originally derived” (Yeo, 2003, p.63). Hence readers were assumed to be familiar with the creation of a common-place book, as active note takers and familiar with the activity of using the alphabet to find something for the purpose of reference.

9.2.4 Author

Chambers’ authorship of the *Cyclopaedia* is unique. First, he took it upon himself to answer how to summarize and organize information as to alleviate the reader from the vast stacks of libraries. While being a solitary author was nothing new, he did so by denying himself the traditional concept of authority. He describes his role in the making of the encyclopaedia through such activities as compiling and summarizing contents, not writing them. In the preface, he remarks that it is “vain to pretend anything of property in things of this nature” (Chambers, 1728, p.xxiv). As such, it is more apt to describe his authorship of the encyclopaedia by how he organized it and not by what he wrote.

Chambers’ *Cyclopaedia* continues to be praised for his innovation of cross-references for which it received much fanfare during the eighteenth century. Perhaps as the greatest form of flattery, it was the model of the famous *Encyclopédie* and Abraham Rees’ *New Cyclopaedia*. As well, it was both glorified and vilified by the editors of the *Encyclopaedia Britannica*. All in all, his work set the stage for the flurry of encyclopaedic activity that was about to take place.

9.3 Diderot and d’Alembert’s *Encyclopédie*, 1751–72 CE

9.3.1 Lexicon

The *Encyclopédie* (1751–72) was edited by two French men, Denis Diderot and Jean d’Alembert. In terms of the purpose of the encyclopaedia, Diderot wrote that “the aim of an *encyclopaedia* is to collect all the knowledge that now lies scattered over the face of the earth, to make known

common phrases, translations and corrections are just a few kinds of marginalia that a reader may make. Jackson’s *Marginalia* (2001) describes this practice in great detail.

its general structure to the men among whom we live, and to transmit it to those who will come after us” (Withers, 1996, p.284). To achieve this goal Diderot & d’Alembert adopted the epistemology of the English philosopher Francis Bacon (1561–1626) for their organizational model. “If we emerge from this vast operation,” Diderot wrote in the Prospectus “our principal debt will be to the chancellor Bacon, who sketched the plan of a universal dictionary of sciences and arts at a time when there were not, so to say, either arts or sciences” (Morely, 1884, p.80). Bacon’s philosophy was established as the foundation for most modern encyclopaedias, in his book *Instauratio Magna* (1620). In this work he divides human understanding into three main forms: reason, memory, and imagination. Stemming from the three categories, the *Encyclopédie* covers many branches of science, as the figure of “The System of Human Knowledge” (*Système figuré des connoissances humaines*) shows (See Appendix, p.163, Figure 2).

However, as much as Diderot was indebted to Bacon, his main concern was not just the rationality of Bacon’s classification, but “the imperative of collecting human knowledge in some manageable form” (Yeo, 1991, p.29). Just like the majority of its predecessors, the *Encyclopédie* was more than a summary of knowledge or a work of reference. Diderot’s intentions were in fact “‘to change the general way of thinking’—not only to challenge traditional religious thought, but to challenge authority [...] in the name of the ‘new science’” (Durbin, 1996, p.127). Considering this, Withers alludes that encyclopaedists had the intention of reframing “[t]raditional learning” as inferior to their new rational form. Thus, “[b]eneath and within the text of the *Encyclopédie* there is, then, a major epistemological shift intended to transform the topography of everything known” (Withers, 1996, p.284). Like Chambers’ claim that his work was integral to the progress of science, Diderot understood that it was the duty of encyclopaedias to aid the development of science. Withers argues that these encyclopaedias “were situated metaphorically in enlightenment questions of order and culturally as texts seeking to direct useful knowledge to social ends” and to bring “knowledge and people together” (1996, p.293). The difference that made the *Encyclopédie* an important work was that it offered new answers (provided by Bacon) to such questions about knowledge and new reasons for such a project.

Beyond the conceptual choice to use Bacon’s philosophical system and promoting science, the *Encyclopédie* also included technical and craft knowledge. Considering that most encyclopaedias of the time concentrated on previous scholarship, Diderot included “knowledge from

common tradespeople” (Wright, 2008, p.148).⁷⁴ Yeo reiterates this fact when he recounts that d’Alembert “believed that the function of the *Encyclopédie* was to contain what was known collectively by the human race” (2007, p.53). This position is manifest in the parts of the “System of Human Knowledge” (Figure 2) which went beyond the trivium and quadrivium of medieval scholars. For instance, *History / Natural / Uses of nature* branches off into *Arts, Crafts, Manufactures* which includes such topics as locksmithing, glass making, cloth-making, tanning, and masonry. Furthermore, what was considered to be respected knowledge was expanded beyond the liberal arts.

Like Chambers, the *Encyclopédie* employed the use of cross-references. In order to provide the reader with a “collection of knowledge in an accessible form” Diderot offered two functions for cross-references (Yeo, 1991, p.27). The first was to “‘elaborate the connections that each special branch of knowledge has with the parent tree’ thereby reinforcing the ‘unity’ of the ‘whole encyclopaedia,’ and second, to ‘highlight contrasts between theories and question tacit assumptions’” (Yeo, 2007, p.56). The first function is obviously used to reinforce the concept of unity. The function of the second is a little less obvious. The cross-references that Diderot uses are not only used to expand scientific concepts but also to be politically subversive.⁷⁵

9.3.2 Connected Literature

As has been stated earlier, the *Encyclopédie* was modelled after Chambers’ *Cyclopaedia* and is in debt to him for the invention of cross-references. In fact, the *Encyclopédie* was originally intended to merely be a translation of the *Cyclopaedia* (Kennedy, 2009, p.34). Distinguishably, Diderot & d’Alembert took the project much further (Loveland, 2006, p.62) with the number of volumes that they added. Also, while Chambers did not always cite his sources, Diderot & d’Alembert saw the connections made between text and source just as important as cross-references. Ferris explains that the *Encyclopédie* not only affirmed “the interrelation of knowledge” but also “account[ed] for its origin” (2003, p.1260). This meant that it became

⁷⁴At the time, there was a perception that dictionaries and encyclopaedias differed in subject matter. While encyclopaedias held to laying out systematic knowledge through the arts and sciences, dictionaries tended to concentrate on geography, history, and biography (Loveland & Kafker, 2009, p.80).

⁷⁵For example, the article ‘Eucharist’ ended with the recommendation “see cannibals” which was a tongue-in-cheek jab at the Catholic church. (Burke, 2000, p.186). This use of an encyclopaedia as an explicit political tool has been noted for being a novel intention in terms of the history of encyclopaedism (Burke, 2000, p.29).

editorial policy to cite all sources used.⁷⁶

9.3.3 Intended Audiences

The editors of the *Encyclopédie* took it upon themselves to make the work more accessible to its readers. D'Alembert describes that the encyclopaedia was meant for nonspecialists. Because of this, he claimed that treatises (long essays about a subject) were not appropriate for the audience (Loveland, 2006, p.63). While they considered their readers as nonspecialists, that does not mean that they did not need their participation. On the economical front, the *Encyclopédie* was sustained through a subscription model. That is to say, that parts of the encyclopaedia were published in instalments and the readers would pay as the encyclopaedia was released. Kennedy describes that given the time between volumes, the encyclopaedists could “solicit and incorporate contributions from the reading public” (2009, p.236). The use of subscription also allowed other forms of interactions between readers and the encyclopaedists. Featherstone & Venn mention that the temporality of the encyclopaedia allowed it to not only “update people on some new technique, scientific discovery or newly discovered and classified flora or fauna” but also to “stimulate public discussion and dialogue” (2006, p.6). These scholars point to the fact that the audience was not thought to be passive readers but were active participants in the creation and economic maintenance of the encyclopaedia.

9.3.4 Authors

One of the attributes that differentiates the *Encyclopédie* from its predecessors is that the content was not only from previous sources but also from 140 living contributors (Kafker, 1973, p.452).⁷⁷ Sullivan states that Diderot & d'Alembert “recruited a wide network of contributors” (1990, p.324). Often one encyclopaedist would ask others to contribute; non-contributors would also attract new contributors. Some contributors volunteered information without request while others “came forward because they wanted to contribute to a work they admired” (Kafker, 1973, p.457). Many of the contributors have been noted to have met Diderot & d'Alembert at coffee houses around Paris (Burke, 2000, p.48). The collaboration described by this organization

⁷⁶Up until this time, the function of citation was not consistent in scholarly works (exemplified by Chambers' qualification that his sources were in the public domain).

⁷⁷Katz notes that the *Encyclopédie* used around 160 contributors (1998, p.36), while Burke reports 135 (2000, p.172)

of contributors was part of Diderot's scheme for the encyclopaedia. As Featherstone & Venn note "Diderot envisaged a voluntary association of individuals responsible for the production of the encyclopaedia, which also suggests a wide public readership" (2006, p.6). While many contributions were voluntary, Diderot & d'Alembert were paid for their work as editors. In the creation of the *Encyclopédie*, Diderot was paid 7,200 livres while d'Alembert received a third of that amount (Burke, 2000, p.172). Given both the social and the economic conditions, both editors were financially invested and intellectually committed to gathering knowledge from a wealth of new sources.

Inevitably, Diderot found that working with living contributors was not without its issues. Kafker begins by stating that in 1768, Diderot expressed that in "addition to some excellent people, there were others who were weak, mediocre, and totally incompetent" (Kafker, 1973, p.452). Further to this point, Diderot wrote in *The Editor's Lament*: "What a mess of egos to manage and disagreements to mediate! And, among my collaborators, what unwillingness there is to get going and be guided by a common discipline!" (Sullivan, 1990, p.325). Kafker contends that such a menagerie of contributors was not a folly, despite Diderot's complaints, but was the main reason that the encyclopaedia was such a success (1973, p.458). Although there was a general unevenness in the writing, it is difficult to imagine that Diderot & d'Alembert could have attracted the number of experts by themselves without the help of their contributors (Kafker, 1973, p.459). Kafker explains that "it was desirable that persons with varied backgrounds and talents select contributors to such a comprehensive work" (1973, p.459). In fact, they were able to include such names as the philosopher Rousseau and the author Voltaire to their list of contributors.

The intersection of using living contributors and the subscription model for their volumes challenged the traditional view that knowledge about the world was static. As Burke notes, the research by living contributors "was connected with the idea that the stock of knowledge was not constant in quality or quantity but could be 'advanced' or 'improved'" (Burke, 2000, p.46). In order to offer a consistent work despite the variety of voices, d'Alembert espouses that the encyclopaedia must "gather them under the same point of view" which was pursued with a great deal of difficulty (Ferris, 2003, p.1260). To this point, d'Alembert reveals that both he and Diderot worked like the philosopher who is placed "above this vast labyrinth at a greatly

elevated point of view from where he may be able to perceive the principal arts and sciences at the same time” (Ferris, 2003, p.1261). While like Chambers, Diderot & d’Alembert saw themselves as overseers of the knowledge collected, they differed in that they also took on the roles as overseers of the entire enterprise, including a large company of contributors. In other words, we see that Diderot & d’Alembert’s enterprise was both innovative and exemplar of the Enlightenment project. By using the philosophy of Bacon as a keystone, they were able to benefit from a variety of contributors who provided new and unique information that made them desirable to the growing literate public.

9.4 The *Encyclopaedia Britannica*, 1771– CE

Out of the major encyclopaedias that started during the eighteenth century, only the *Encyclopaedia Britannica* (1771–) has continued to the present day.⁷⁸ From its meagre beginnings as a three-volume set in 1771 to its most voluminous thirty-five in 1926; from its hot-metal printing, to phototype, to its digital existence on the web, the *Britannica* has seen two hundred and forty years of change. Started by William Smellie in 1768, it began as a rather rudimentary encyclopaedia. Despite the Preface’s boastful claim that it was “superior to all previous encyclopaedias”, the three volumes of the First Edition paled in comparison to its competitors, the *Cyclopaedia* and the *Encyclopédie* (Kafker & Loveland, p.299). Articles of the First Edition were little more than a list of headings adorned with typographical errors. Its subsequent editions have continually corrected, adapted, and built upon the original to create a work that has been revered for over two centuries.

9.4.1 Lexicon

The organizational structure of the first edition of the *Britannica* tends to be described in two ways. While it is alphabetically ordered with short summary articles tied together through cross-references, there are also treatises that explain the connection between major subjects. Various editors of the *Britannica* see the use of treatises in different lights. Some see treatises as a necessity to relate the proper order of subjects while others see them as obstacles in referencing information. An example of the former position comes from the editor of the Third Edition who

⁷⁸Albeit, *Britannica*’s last printed version is the 2010 edition.

asserts that his edition paid special attention to classifying “objects in human systems as they are in fact classed in the system of nature” (EB, 1790b, p.ii). This assumption was revisited in the Seventh Edition, where editor Macvey Napier states that the “sciences cannot be taught or understood without being viewed continuously, in their natural state of unity and coherency” (1842, v.1, p.ix). Justifiably, even these editors have some doubt as to the existence of a truly “natural” order to the sciences.

After reviewing several philosophers’ categories of nature the editor of the 1790 preface admits, with resignation, that in view of all of the great minds and their different classifications, the attempt to come to a definitive set of categories will never be accomplished (EB, 1790b, p.ii). A similar survey is conducted by the editor Dugald Stewart in 1853. He describes how he initially planned to take the outline conceived by Diderot’s adaptation of Francis Bacon’s “general survey” of human knowledge and adapt it to “the present advanced state of the sciences” (Stewart, 1853, p.1). Like his 1790 predecessor, he found flaws in such a project and had to abandon it. He cautions that not only was the “division of faculties” of Bacon, and later d’Alembert and Diderot, unsatisfactory “but that every future attempt of the same kind may be expected to be liable to similar objections” (Stewart, 1853, p.6).⁷⁹⁸⁰

The 1967 Preface argues that definitive categorization of the sciences is so difficult because as knowledge increases, it also “changes its shape and character. As interests change, and more and more workers devote their efforts more or less exclusively to increasingly specific areas of knowledge, new specialities split off from old ones; new forms of expertise develop; whole new sciences are born” (Preece, 1967, p.vii). Despite the fragmentation editor Warren Preece foresaw, he held the conviction that knowledge is even “more subtly interrelated” (1967, p.vii), that fragmentation is actually evidence of a greater unity of knowledge. The editors rectify the issue of fragmentation by acknowledging that “although a general encyclopaedia takes all of knowledge as its province, no encyclopaedia can cover all of knowledge” (due to the physical

⁷⁹Stewart’s final decision was to break the world’s knowledge into the two most basic elements, matter and mind, for which no “branch of human knowledge, no work of human skill, can be mentioned, which does not obviously fall under” (1853, p. 10). Given these two main categories the Eighth edition contains a number of treatises that show how each discipline had progressed to its current state (Stewart, 1853, p.10).

⁸⁰One of the interesting comments that Stewart makes about describing the progress of all the particular sciences is that while the “discoveries of the Mathematician or of the Chemist” show unprecedented growth “improvements in Moral and Political Science do not strike the imagination with nearly so great force” (1853, p.11). This admission belies that perhaps describing the subjects in terms of progress may not have been the best strategy since some subjects may not be progressive in nature.

nature of the book) and thus “[i]t is the editor’s responsibility to decide,[...] what can go into a set of books” (Preece, 1967, p.vii).

The 1984 Preface reiterates this position when editor Robert Hutchins states that while many encyclopaedias have argued utility in terms of organization “the fact is that the argument misses the point” (1984, v.1, p.x). He explains that knowledge is not organized purely by “either a topical or an alphabetical form”, but can only be understood as a synthesis of the two (Hutchins, 1984, v.1, p.x). More importantly, Hutchins insists that encyclopaedias need to question the degree to which an “organizing principle” presents knowledge through synthesis or fragmentation. This dichotomy was the starting point in 1947 for developing a new and radically different Fifteenth edition. While the *Britannica* had performed the function of reference (alphabetical organization) successfully in past decades, it had not acted as an adequate “instrument of education” (Hutchins, 1984, p.v). He found that if topics are used to reveal connections between knowledge (as in the case of an education), it is difficult to find specific information. Alternatively, an alphabetically organized encyclopaedia is easily navigated but fails to provide an adequate education through the connections between subjects. Correspondingly, the *Britannica* was redesigned to be both a work of reference and an instrument of education that was divided into three divisions, each one constructed to address a particular way that readers use an encyclopaedia (Hutchins, 1984, p.v). The first section is the *Propaedia*, an introductory series of treatises that outlines the traditional fields of knowledge. Durbin defines the *Propaedia* as a sequential framework “that enables you to learn on your own, at your own pace, using the pages of *Britannica* as your personal tutors”, and subsequently fulfilling the educational role (Durbin, 1996, p.131). Following this set is the *Micropaedia*, which is an index-like reference section that offers micro summaries of an article and the cross-references. Finally comes the *Macropaedia* which closely resembles the summary articles that *Britannica* is known for.

Overall, the emphasis within all these changes to the encyclopaedia were directed at the idea that the encyclopaedia must be useful to its readers, and as the 1984 Preface declares “utility is in fact the name of the unit in which encyclopaedias must inevitably be measured” (Hutchins, 1984, p.xi). Lamentably, it must have been with great disappointment that the Fifteenth edition was poorly received by the public. The majority of the complaints were not about the quality of the content but the difficulty in using it. The *Atlantic Monthly* ran a scathing review with

the cheeky title “*Britannica 3, Failures of*” (Wolff, 1976). It had turned out that after years of refining its reference function, people had little use for the three “-paedia” divisions. The next revision in 1984 responded to the feedback of its readers and included a straight-forward index. The final print edition in 2010 continued to divide the encyclopaedia into these three -paedia sets with the additional index.

In 1994, the *Britannica* began producing a CD-ROM version of the encyclopaedia. This digital edition included not only images, diagrams, and words, but sounds and videos as well. This format continues to be produced today as a DVD. In keeping up with technological developments, *Britannica* also launched a website in 1994 and is today the 7031th most visited site on the World Wide Web (Alexa.com, 2012b). This recent history shows that the *Britannica* has been committed to adapting the encyclopaedia to new technologies.⁸¹

Beyond the organizational structure of the *Britannica*, it has also been organized by the information that it includes. The editor of the 1910 ‘Editorial Introduction’ describes the substance of the encyclopaedia as being dominated by the “historical point of view” because “what *has been*, if accurately recorded, remains permanently true as such” (Chisholm, 1910, pp.xix).⁸² Though he admits that anything recorded close to the publication date may “become inadequate almost as soon as it is in the hands of the reader” (Macvey, 1842, pp.xix). In a similar sense, the inclusion of biography was, until 1902, relegated to only deceased personalities (Einbender, 1962, p.116). Beyond the presence of a historical view of encyclopaedic knowledge, Macvey explains that the *Britannica* relies heavily on “the use of statistics” “the spirit of science” “the comparative method”, “the objective view”, and “the art of illustration” (1842, pp.xix-xxi).

In view of the “spirit of science”, Macvey describes that the “object of the present work is to furnish accounts of all subjects which shall really explain their meaning to those who desire accurate information” (1842, p.xxi). Hence, when it comes to divergent opinions, the encyclopaedia recognizes these differences by “stating a belief in the terms, and according to the

⁸¹Despite such changes, the *Britannica* has lost its long-held popularity. This development will be further discussed in the conclusion of the literature review.

⁸²Interestingly, the 1842 edition included both history and political science, subjects that are surprisingly considered unencyclopaedic at the time (Macvey, 1842, p.xi, p.xxxvi). By some critics of the *Britannica* “To include history, in any form, was considered [...] an abuse of the purpose of an encyclopaedia” (Macvey, 1842, p.xv). Understandably, the editors believed it was important to include both and thus politics was first published as a science in Britain in the 1801 Supplement of the *Britannica* (Macvey, 1842, p.xxvi) while history continues to see extensive use throughout the editions of the *Britannica*.

interpretation, accepted by those who hold it” (Macvey, 1842, p.xxi). He continues to explain that in the case of dogmatic subjects “contributors of all shades of opinion have co-operated in the work” (Macvey, 1842, p.xxi) and that these perspectives are represented only through the “ultimate criterion” of consensus (Macvey, 1842, p.xxi). However, despite the claim for all kinds of contributors, the editor reveals that such consensus is garnered only from “countries which represent Western civilization” (Macvey, 1842, p.xxi). Therefore, the spirit of science is fundamentally a Western idea realized only through Western contributors authorized by the *Britannica*.

The 1842 edition of the *Britannica* was the first edition to accompany articles with statistics. As stated in the 1910 Preface, statistics play a pivotal role in representing encyclopaedic knowledge. “Statistics are used here as an illustration of the substantial existing conditions and of real process” (Chisholm, 1910, p.xix). He also warns that the statistics used are not to display the most up-to-date information but “to show the representative conditions of abiding importance” (Chisholm, 1910, p.xix) as well as an expression of “the substantial results of human activity” (Chisholm, 1910, p.xx).⁸³ Other illustrative devices that the *Britannica* has used include maps (since 1790), illustrations (since 1878), and photography (since 1910).

Editor of the 1878 Preface, T.S. Baynes, expounds how the *Britannica* is not only a work of reference but is “to some extent at least, an instrument as well as a register of scientific progress” (1878, p.v). This was achieved through the inclusion of articles that were created through “the fruits of original observation and research” (Baynes, 1878 p.v). Therefore the 1878 editors envisioned the *Britannica* being a pivotal cog in the machine of science. According to one historian, the *Britannica* did indeed function in this capacity: as both a “vehicle for the dissemination of science to the general public” as well as “a medium [...] for the presentation of original research” (Yeo, 1991, p.46). Interestingly, Baynes provides a warning. He warns that although the new directions taken in science are reflected in the *Britannica*, they “must be regarded [...] as tentative and provisional” (1878 p.vi). To add to this, Baynes advocates that it is not in the interest of the *Britannica* to involve itself in opinion. Instead it deals with knowledge that has been dealt with critically and historically (Baynes, 1878, p.viii). This statement is at

⁸³This kind of empirical evidence of human activity is considered by Bazerman as having preferential treatment “within the values, norms, expectations, procedures, and evaluations of the scientific community” (1988, p.190).

odds with his 1842 predecessor who states that the encyclopaedia is the domain of “general subjects [that] are discussed in systematic or complete treatises—including facts, opinions, and authorities” (Macvey, 1842, p.xl). Baynes continues by stating that the duty of the *Britannica* “is to give an accurate account of the facts and an impartial summary of results in every department of inquiry and research” (1878, p.vii). That being said, opinions and bias sometimes replace facts and have been the subject of many criticisms.⁸⁴

Baynes’ greatest reason for providing original scientific research is relayed in the following quotation: “The available facts of human history, collected over the widest areas, are carefully co-ordinated and grouped together, in the hope of ultimately evolving the laws of progress, moral and material, which underlie them, and which, when evolved, will help to connect and interpret the whole onward movement of the race” (1878, p.vii). However, by the time the 1967 edition was published it ceased to include original research as the encyclopaedia was reconstituted as merely a “report on the state of knowledge” (Preece, 1967, p.vii). This attitude has remained to the present day. Thus there has been a real change in the function of the *Britannica* as it had once documented the advances in science as they happened, it now only reports on stable and authorized facts. Since then, the substance of the encyclopaedia has been revisited in the 1984 preface by describing six encyclopaedic qualities. Hutchins describes that the encyclopaedia “must speak from and with authority”; must cover the “whole circle of learning”; must be summary in nature; be accessible through the structure of organization of content; be accurate; and be internationally composed (1984, p.xi). This list of attributes (with the exception of the last item) accurately describes the subject matter that the *Britannica* has contained for the past two hundred and forty years.

9.4.2 Connected Literature

Many of the prefaces of the *Britannica* connect themselves to past and rival encyclopaedias by arguing the virtue of their arrangement. In Genette’s term, this is a form of metatextual commentary. The 1771 Preface argues that the organization of the *Cyclopaedia*, dismembered the Sciences in contrast to its own treatises that “digested the principles of every science”,

⁸⁴See Einbender (1964), and Thomas (1992).

providing a more acceptable form (EB, 1771b, p.v).⁸⁵ Its self-confidence continues by stating that it spares “readers the ‘labor of hunting for science through [...] a labyrinth’” (Loveland, 2006, p.57). This final comment is directed toward the purely alphabetical organization and usage of cross-references used by the *Cyclopaedia*.

In the Second Edition of the *Britannica* (1778b), the *Cyclopaedia* is again called to answer for its organizational decisions. The editor proclaims “when topics, far from being digested into a system, or disposed in their natural order, are violently dilacerated, and, without any regard to their proper positions, huddled together as the order of the letters which constitute their technical terms determine, such a work should rather be called a book of shreds and patches” and “rather a chaos of detached and heterogenous ideas, than a regular intellect” (EB, 1778b, p.iv). As this quotation demonstrates, there is a strong belief that knowledge must be correctly ordered or else it would be detrimental to the reader. In the Seventh Edition (1842), Macvey attacks the *Cyclopaedia*, but this time for not showing the scientific processes that produced the conclusions that were reported (1842, p.ix).

Diderot & d’Alembert’s *Encyclopédie* is also reproached by the *Britannica*’s frontmatter. Dugland Stewart’s preface for the Eighth Edition (1853) asserts that “it cannot be denied that d’Alembert has displayed much ingenuity and invention; but the depth and the solidity of his general train of thought may be questioned” (1853, p.2). The commentary that these editors provide demonstrates both a kinship with past encyclopaedias as well as the belief that organization of information is one of the main distinguishing attributes of encyclopaedias.

The practice of citation (or intertextuality in Genette’s terms) has changed throughout the editions of the *Britannica*.⁸⁶ The second edition creates references to works through short marginal citations. For example, the 1778, 1790, and 1798 editions of the article ‘America’ cite only the title and volume of a work to create a reference. This practice changes in the 1842 edition where references are recorded in full (author, title, city, year, page) and moved to the foot of the page in the form of footnotes. This practice continues until 1910 when references to other works are given their own heading at the end of an author’s section. Sometimes these headings are entitled ‘Authorities’ or ‘Bibliography’. By 1968, this practice is systematized

⁸⁵ As mentioned previously, Chambers argues in his preface that there is no natural way to discuss the branches of sciences.

⁸⁶ These changes in citation practices have been observed through original research.

by consolidating the bibliographic notes of each author into notes at the end of an article. While topical divisions are made within the bibliography these headings do not always directly correspond to the topics used within the article itself. The final change to citation practices occurs on the current web edition of the encyclopaedia. External works are categorized in a number of ways depending on their use. If they are direct citations, the section 'Citations' lists them. External printed information is listed in "References" while online sources are listed in 'Related Links'. The change of these formal matters reflects that the value of the source of material has changed throughout the *Britannica's* history. Such changes can be best described as a continuing interest and revision of how encyclopaedic knowledge is presented and validated.

9.4.3 Intended Audiences

As discussed earlier, the description of what a reader did with the encyclopaedia was subject to revision over the past two hundred and forty years. In the Second Edition, three kinds of readers are identified. The first is the "systematic reader" who is "fully and regularly informed by turning to the general name of the science which he wishes to explore" (EB, 1778b, p.iv). The second kind of reader, the specialist, is one who is "already acquainted with the whole, wishes only to consult particular topics" (EB, 1778b, p.iv). The envisioned specialist reader can be seen in the type of language that different editions used. The Eleventh Edition commonly employed extensive footnotes that contained Greek and Latin quotations. Einbinder notes that this "approach was fashionable in the eleventh edition because it was addressed to an educated elite who had been exposed to the Greek and Latin in British public schools" (1964, p.153). Knowledge of these two languages assumed that the readers had a certain level and experience of education. Be that as it may, if the Eleventh edition was intended for specialists, it was done so at the expense of the understanding of laypeople (Einbender, 1964, p.152).

Interestingly, catering to both types of readers lead to internal struggles within the corporation. During the time that the chief editor Macvey Napier was in charge (1830–1842) he was requested by the publisher to reduce the overall size of the encyclopaedia to be more economical. Napier fought back by making the case that reducing the size would result in making a choice between either diminishing "the quantity of miscellaneous matter more particularly adapted to the wants and tastes of ordinary readers" or "treat important subjects in a way too curt and

superficial to satisfy those of a higher class” (Kogan, 1958, p.44). As such, both tactics would inevitably lower the *Britannica*’s “popularity and reputation, and enable its rivals to gain an ascendancy at its expense” (Kogan, 1958, p.44). Napier’s argument won out and the Seventh edition was substantially longer than the previous edition.

Finally, the third kind of readers “are willing to content themselves with partial and detached views of things” (EB, 1778b, p.iv).⁸⁷ These three archetypes are moreover used as a justification for how the information is displayed.⁸⁸ In describing these three kinds of readers, the *Encyclopaedia Britannica* aimed to enable “any man of ordinary parts” to “learn the principles” of the many disciplines of knowledge (Loveland, 2006, p.67). This attitude is present in the *Universal History of Arts and Sciences* which claimed that readers could come to “perfect” knowledge “without the assistance of a master” (Loveland, 2006, p.70). Charles van Doren, an editor for the *Britannica*, agrees with this point when he insists that any encyclopaedia is “an instrument of enlightenment, and an ideal encyclopaedia would be an educational institution of the highest importance, with a far-reaching effect” (van Doren, 1962, p.26). This educational purpose of encyclopaedias to assist students has been reiterated in the Prefaces of the Eleventh, Fourteenth, Fifteenth, and online editions of the *Britannica*. Each of these editions claim that developing content for students is pivotal to the creation of encyclopaedias. Therefore beyond the function of utility and reference, the *Britannica* also incorporates an educational function.

By 1967, the concept of the reader had taken a slightly new direction. The preface advocates that an editor must be in agreement with the “user” (Preece, 1967, p.vii). From this point on, the concept of the reader as specialist came to hold less precedence while the reader as user became one of great importance. The 1967 edition outlines its objectives of editorial policy to create an encyclopaedia that was readable and intelligible “to the curious, intelligent layman”, provides information that is integrated and coherent, limiting the amount of duplication of information, written objectively and with neutrality in mind, and is a combination topical and alphabetical

⁸⁷Einbinder comments that while some readers use the *Britannica* as a reference work, for others “it is a prestige symbol or tangible sign of intellectual sophistication; while for still others, it is a testimonial to the persuasive skill of a salesman” (1964, p.161).

⁸⁸In a clever ploy to avoid responsibility for the inexact nature of their system of organization, the 1790 editors place the blame of missing information on the shoulders of the readers: “it must [...] be acknowledged that such omissions are neither numerous nor very important; for many subjects, which have been supported to be admitted, are treated under titles different from those under which they have been looked for” (EB, 1790b, p.vii). So despite their extreme interest in the utility of their work, the editors shift the blame of any problems with their system.

arrangement (Preece, 1967, p.xv). This edition did away with treatises on specific subjects and presented everything in a strictly alphabetical fashion to increase the accessibility of the content for the reader (Preece, 1967, p.viii). However, this position would be reversed in 1974 with the creation of the Fifteenth edition of the *Britannica*.

The continuation of the 1974 plan of three -paedias in 1984 states that a user in need of the *Propapaedia* would be “interested in an understanding of the objects of studies of all of the sciences as they relate to something grander than the disciplines themselves” (Hutchins, 1984, p.xi). Hutchins instructs that the user in need of the *Micropaedia* would be looking up something that was limited in scope (1984, p.xi). Finally, the user of the *Macropaedia* would be someone looking for “information about a broader, but still relatively limited, subject” (Hutchins, 1984, p.xi). These three descriptions of readers were seen as the rationale for the monumental change in the *Britannica*’s structure. At the same time as these readers were articulated, the idea of the “specialist” reader was considered to be ill-fitted to encyclopaedic knowledge. The Fifteenth edition resigned that the *Britannica* could not offer the “specialist” reader anything of worth. In fact, there was doubt whether any edition of any encyclopaedia ever had (Hutchins, 1984, p.xii). Thus as the envisioned reader has changed, so to has the organization of the encyclopaedia.

Audiences have also been considered as pivotal aspects of the business models for encyclopaedic publications. Editors of the early *Britannica* often described their audience as subscribers or commercial partners in the corporate enterprise of encyclopaedia publication. As it was with the *Cyclopaedia* and *Encyclopédie*, the *Britannica* used a subscription model to finance its project. It was common among the earlier editions to include a list of subscribers in the frontmatter. Even though this is evidence of early audience interaction, Burke describes how “[s]ubscription lists are also a vivid reminder of the problem of the limitations to individual access to knowledge at this time” (2000, p.177). He explains that as the short lists of subscribers indicates, only a fraction of the population could afford to purchase encyclopaedias. In time libraries would buy entire sets for their community and access to the encyclopaedias would eventually increase. The audience as subscribers continues to the current web version where the term “members” is used to identify people who pay for access to the online *Britannica*.

Over the past two hundred and forty years the *Britannica* has conceptualized its audience in a variety of ways. During the eighteenth and nineteenth centuries, readers were often thought to

be specialists who used the articles to become aware of the latest developments in their science. However, this conceptualization was later dropped and so the inclusion of original research was no longer deemed suitable as content for the encyclopaedia. Long standing audience types such as the student and the average reader have been prevalent since the beginning of the *Britannica* and continue to today. In both cases, there is the understanding that an encyclopaedia can be used to educate its reader on the proper connections between subjects. Finally, the intended audience has been subscribers to the encyclopaedia, individuals who have provided the capital necessary to finance the continual production of such an enormous project.

9.4.4 Authors

The authors of the *Britannica* have played different roles in the development of the encyclopaedia. Notably, the First Edition “was the first encyclopaedia to maintain a permanent staff of editors and sub-editors to facilitate a constant process of revision and renewal” (Doody, 2009, p.20). This points out two important developments that must be unpacked.

The first is that by the time the *Britannica* was initially published, the creation of an encyclopaedia had become a viable commercial enterprise. While a permanent staff of editors was employed to compile and organize previous sources, their contributions were done so anonymously until 1842.⁸⁹ While the Second Edition merely “collected and compared” the best authors (EB, 1778b, p.v), the 1878 edition employed scientific experts to write articles with original research and to grant an edge over competitors (Yeo, 1991, p.46).⁹⁰ Einbender argues that hiring experts was a “natural” decision for encyclopaedias to create an “authoritative account of particular subjects”. But, as he points out, experts tend to be so intimately familiar with the subject that encyclopaedic discourse tends to emphasize facts at the expense of ideas (Einbender, 1964, p.151). Regardless, the inclusion of such experts was seen as a selling point for an 1858 advertisement that explains how the “wonderful story of the 19th Century is told

⁸⁹Their authorship can only be found on the title page of the first edition (1771) of the *Encyclopaedia Britannica* as being written “by a society of gentlemen in Scotland”, after which there is a list of sources of compiled writing that are used in the encyclopaedia. The next two editions repeat this practice of distinguishing between the anonymous editors and the dead authors that are used as sources. It is not until the 1842 edition that a list of living contributors is first displayed.

⁹⁰Loveland notes that in 1701 “the biblical scholar Richard Simon anticipated the modern encyclopaedia by arguing for the participation of at least twenty specialists” (2006, p.60). It is curious as to how he came up with that exact number.

by the men who made its greatness” (Kogan, 1958, p.82). To this end, paid experts became an essential aspect of the *Britannica*’s business model. Surprisingly, payment for these authors was sometimes quite low. Between 1929 and 1964 contributors of the *Britannica* were paid two cents a word.⁹¹ Therefore, while the experts have been paid by the *Britannica* for their work, the encyclopaedia relied heavily on contributor’s altruistic beliefs (in being part of such a prestigious intellectual project) to get away with relatively low pay.

The second part of Doody’s quotation, “to facilitate a constant process of revision and renewal” indicates that the new encyclopaedia needed to be continuously updated (2009, p.20). Epistemologically speaking, the necessity of updating the encyclopaedia is the physical embodiment of perceiving knowledge as infinitely changing. In this regard, all of the editions up to the Thirteenth followed the same publication cycle; creating a staff of editors and writers, putting out a new edition, and then trying to keep it in print until its content was out-of-date. Then the whole process would start again (Ashmore, 1962, p.15). Accordingly, questions of “datedness” became a real concern for encyclopaedists. Knowledge had to be relevant now, to the people, and be organized in the best way to provide access. In all, these encyclopaedias manifested the progress of science, the speed of industry, and the growing sense of a world expanding before their very eyes. In 1929, the *Britannica* moved toward a new model of publication that would provide current reports on the state of knowledge. Given the advancements in printing the *Britannica* began its policy of “continuous revision” which directed the annual update of all articles in need of adjustments (Preece, 1984, Preface, p.ix).⁹² The encyclopaedia was then augmented by the annual *Year Book*, which would update statistical and geographic information that was prone to change from year to year. To ensure that the 1967 edition was in keeping with society-at-large, it took great lengths to ensure that its contributors for each article included at least one international expert (not from the United States) (Preece, 1967, p.xvi). This strategy was a response to guarantee a greater fidelity in terms of facts and truth. Indeed, the 1967 edition had grown to the size of 4300–4500 editors and contributors, many of them being internationally

⁹¹ As Einbinder notes, Albert Einstein wrote his article ‘Space-Time’ for only \$86.50 (1964, p.266). At the rate of two cents a word, the Fourteenth Edition paid an estimated \$760,000 to contributors. To put this amount in perspective, the 1960 advertising budget of the *Britannica* was over five times more, nearly \$4,000,000 (Einbinder, 1964, p.269).

⁹² While the policy was intended to cure many temporal flaws, it oftentimes fell flat. Einbinder notes that an essay on Greek Art from 1963 was originally published in 1910, half a century earlier. Similarly, an article about the Punic War from the same edition contained “recent research” from 1901 and 1902 (1962, p.280).

renowned scholars (Preece, 1967, p.xviii). The 1984 edition also notes its policy for articles to be composed by groups of international writers, allowing for only one American writer per article. Hutchins explains that while “the new Britannica is uniquely a product of the culture of the West, it must not be parochially Western in its view of learning” (1984, p.xi). The inclusion of international authors and annual updates were all developed to make the product more useful and indispensable to its readers.

While the activities of editors and contributors are specific to encyclopaedism, the personalities who make up these roles are by no means homogenous. Among their ranks have been scientists, philosophers, scholars, amateurs, experts, clerics, users, novelists, and artists. What must not be forgotten about these different roles is that throughout its history, the majority of *Britannica* editors and contributors have been overwhelmingly male. The first signed women contributors to the *Britannica* were in 1910.⁹³

In brief, the authors of the *Britannica* have fulfilled a great number of roles. The first editors’ role was to compile and summarize previous work and to remain anonymous. It is not until the nineteenth century that authorship was shared between editors and contributors, some of whom were notable authorities or celebrities in their field. Then the twentieth century provided the atmosphere to acknowledge contributions to the encyclopaedia made by women. By the mid-twentieth century, policies were created to ensure that all articles were written by many international hands. The common pattern of these developments is the quest for creating an encyclopaedia that could be deemed objective and unbiased in its representation of facts. This quest has not only opened up new subject areas to explore but has continually breeched the limits of who has been allowed to author an article.

9.5 Wikipedia, 2001– CE

Wikipedia’s success started with the failure of Jimmy Wales’ first attempt at making an online encyclopaedia, *Nupedia* in 2000. Benkler describes that it was built on the traditional production model of using the formal peer-reviewed process for contributions from PhDs (2006, p.70).

⁹³ At a party celebrating the inclusion of such “women of letters”, Hogarth, one of the newly honoured female contributors, noted that while the Eleventh Edition had thirty-five signed women, the Tenth Edition had thirty-seven unsigned female writers (Thomas, 1992, pp.18–19). As Thomas notes, beyond the signed authors, a large amount of women’s work on the Eleventh edition was rendered invisible (1992, p.19). While the inclusion of women was deemed a great movement forward, 98 per cent of the 1910 contributors were men.

The only notable difference was that instead of buying a set of books, the product would be released free to the public via the Internet. What caused the project to fail was a lack of enough “high-quality contributions” to form a coherent encyclopaedia. Subsequently, the few contributions that were written became the starting point of Wikipedia (Benkler, 2006, p.70). What Benkler stresses is that once Wales dropped the “authority” of the writer as a pivotal characteristic, Wikipedia developed “a radically new form of encyclopaedia writing” (2006, p.70). Benkler lauds that what Wikipedia offers, unlike its expert-based competitors, is the opportunity for individuals to participate in the “writing” of their culture and to emphasize topics that are important to them (2006, p.293).

9.5.1 Lexicon

According to Wikipedia:Five pillars (2005a), it describes itself as an “online encyclopaedia” that acts as “a comprehensive written compendium that contains information on all branches of knowledge” (Wikipedia, 2007b). What exactly populates these branches has been left up to the voluntary actions of over twenty-six and half million users to decide (Wikipedia, 2012a). Unlike previous encyclopaedias, the articles of Wikipedia have not been planned from the beginning. Content has been added as people contribute to the project, allowing the structure to grow and adapt depending on the needs of the editors. On the one hand, this structure has formed through the same textual devices that Chambers used to link his encyclopaedia together; through the mixed use of alphabetization and cross-references (or hyperlinks in the case of Wikipedia). On the other hand, Wikipedia is also organized through a hierarchy of topics and a database system.

The first kind of organization, the hypertext, has been the oft glorified textual device that allows one bit of content to be connected dynamically to a whole host of others. Castells defines hypertext as “an actual interactive system, digitally communicated and electronically operated in which all the bits and pieces of cultural expression, present, past and future, in all their manifestations, could coexist and be recombined” (2001, p.202). This quotation conceptualizes hypertext as a device that places “cultural expression” within a fluid and ever-changing environment. Tapscott describes that the “new Web”, empowered by hypertext “is the most robust platform yet for facilitating and accelerating new creative disruptions” (2006, p.31).⁹⁴

⁹⁴For Tapscott, the web of links and the software that supports it disrupts the professional and credentialed

So while Chambers used cross-references as merely a means to circumnavigate the problems of organizing knowledge, hypertext (along with the relatively low cost of access and editing) is a device used to circumnavigate the problems of representing knowledge. Moreover, in the context of Wikipedia, hypertextual links provide the encyclopaedia a great level of flexibility in terms of what information is connected, in addition to the potential meanings that can be interpreted from those connections.

Encapsulating the hypertextual organization of Wikipedia is the database. Lev Manovich explains that database “is defined as a structured collection of data” and the structure can be organized in a variety of ways, albeit hierarchical, network, relational or object-oriented (2000, p.177). Manovich explains that both digital storage media including the Web “proved to be particularly receptive to traditional genres which already had a database-like structure” (2000, p.178). One such traditional genre he is speaking of is encyclopaedias “which are collections by their very definition” (Manovich, 2000, p.177). What the Web has afforded the organization of encyclopaedias (which applies to the *Britannica* as well) is a structure that “is organised for fast search and retrieval” and “which the user can perform various operations: view, navigate, search” (Manovich, 2000, p.177). In addition to helping users find the information they need, electronic databases make it easy to “add new elements to the end of a list as it is to insert them anywhere in it” (Manovich, 2000, p.178), an affordance that print could not offer. Indeed Wikipedia utilizes the database structure created by Ward Cunningham, a database that he claims is “the simplest online database that could possibly work” (Leuf & Cunningham, 2001, p.15).

While databases are collections of information, the activity of making connections between content resides in the agency of its actors. The most common practice of organization is enacted either by individual users or by bots that make inter-article connections. Alternatively, the structure of Wikipedia is also dependant on connections made between categories of articles. As Featherstone & Venn contend, those invested in digital storage and retrieval software and devices, must come face to face with the problematics of the “structure and classification of the world” (2006, p.2). Indeed, some Wikipedians have taken to the problem of classification with great fervour. Editors who participate in the Wikiproject:Outlines argue that “Wikipedia’s goal

authorities that once monopolized the production of both knowledge and culture. Lévy saw the potential for hypertexts in the early 1990s “as a textual machine that could profoundly change writing, and therefore thought” (2011, p.8).

is to provide knowledge. But knowledge is only useful if you can find it. The main limitation of a search box is that it can only look for what you already know exists. Our mission is to show what exists and provide easy access to that knowledge” (Wikipedia, 2007a). Nordloh (1998) agrees with this point by stating that the most prevailing system of organization, the electronic search, is still not quite perfect. He argues that while the algorithms it uses are complex, they do not tell us “what is of value” (p.517). He explains that the “more rapid and chaotic the proliferation of information, then, the more necessary become thoughtful and well-organized reference resources” (Nordloh, 1998, p.517). Some Wikipedians agree with this argument, claiming that the project would fail at providing knowledge if connections between articles relied solely on the myopic connections made by editors who concentrate on isolated articles. Accordingly, these editors of categories have taken it upon themselves to classify the articles manually in order to “help users understand how the category relates to other categories” and “help users jump up to higher levels on the hierarchy” (Wikipedia, 2005b). Further to this point, as early as May 2004, the page Wikipedia:Categorization states that the “central goal of the category system is to provide links to all Wikipedia articles in a hierarchy of categories which readers can browse, knowing essential, defining characteristics of a topic, and quickly find sets of articles on topics that are defined by those characteristics.” (Wikipedia, 2004c). In keeping with hyperlinks, these categories are not mutually exclusive but can link to other categories and articles in a variety of ways, providing a mix of both hierarchical and lateral connections. As such, Wikipedians have gone to great lengths to establish a hierarchy of articles (Figure 3, Appendix, p.164)⁹⁵ The creation of these categories indicates that despite the initial lack of content structure and individual editors making linking decisions, this group of editors see that as an encyclopaedia, Wikipedia must be connected along hierarchical lines if it is to remain useful. As this activity of hierarchical organization shows, there is worth in examining how

⁹⁵The current categories used are first divided among four types of content: ‘encyclopaedic content’, ‘Navigational content’, ‘Wikipedia administration’, ‘Content maintenance and help’ (Wikipedia, 2007c). The very presence of administration, maintenance, and help as top level categories is evidence of the amount of energy that goes into organizing and facilitating the activities of Wikipedians. This has also been noticed by the study Kittur, Suh, Pendleton & Chi (2007) who observe that overall “user, user talk, procedure, and other non-article pages have become a larger percentage of the total edits made” (p.455). The two other categories, encyclopaedic and navigational content, structure the front-end experience for users. Under ‘encyclopaedic content’, it is further divided into articles, featured content, glossaries, lists, and timelines; under ‘Navigational content’: wikibooks, categories, indexes, outlines, and portals. It takes one more level before familiar encyclopaedic topics appear where we see topics as geography, arts, history, mathematics, and science.

Wikipedia has constructed its epistemological beliefs. Unfortunately this aspect of Wikipedia has been largely ignored by epistemologists⁹⁶

In keeping with all the other encyclopaedias, Wikipedia uses summary as a form of writing to describe its knowledge. The rationale of the summary style is explained as to “summarize and distribute information across related articles in a way that can serve readers who want varying amounts of details, thus giving readers the ability to zoom to the level of details they need and not exhausting those who need a primer on a whole topic” (Wikipedia, 2004a). Additionally, summary is used so as to not overwhelm “the reader with too much text up front” and to keep article sizes down to keep a topic “easy to find or read” (Wikipedia, 2004a). Wikipedia also defines that within summarization is the need to categorize the information. In this case, the core elements of an article are summarized through the defining characteristics that secondary sources “commonly and consistently define, in prose, the subject as having” (Wikipedia, 2006a). What is unique about Wikipedia in this regard is that it explicitly details the process and reasons for summarization, an aspect that is relatively absent from the prefaces of the other encyclopaedias.

Beyond the actual connections between articles and the use of summary, Wikipedia is held together by the rules and policies that are abided by Wikipedians. Most notably, Wikipedia has five pillars,⁹⁷ which function as the foundation for coordinating people’s actions. Reagle (2008) observes that the continuing success of Wikipedia relies specifically on the interwoven policies of “verifiability” and “no original research” that stem from the “neutral point of view” (NPOV) pillar (p.11). Benkler confirms that it is an “effort to represent sympathetically all views on a subject, rather than to achieve objectivity” (2006, p.71). Reagle explains that a user can report on the fact of existing opinions, but not use the opinion as a substitute for fact (2008, p.55). To explain, in order for the free and open encyclopaedia to be a reputable source of information, Wikipedians “strive for articles that document and explain the major points of view in a balanced and impartial manner” (Wikipedia, 2005a). They do so by adopting the practice of “citing verifiable, authoritative sources”. An interesting aspect of this policy is that it only

⁹⁶In 2009 the epistemology journal *Episteme* an entire issue was devoted to discussing the epistemology of mass collaboration (Fallis, 2009). However, every article that discussed the epistemology of Wikipedia only addressed the question of whether Wikipedia is a reliable source of knowledge.

⁹⁷The Five Pillars of Wikipedia: 1) Wikipedia is an encyclopaedia. 2) Wikipedia is written from a neutral point of view. 3) Wikipedia is free content that anyone can edit, use, modify, and distribute. 4) Editors should interact with each other in a respectful and civil manner. 5) Wikipedia does not have firm rules (Wikipedia, 2005a).

became an active policy part way through 2003 (Beschastnikh *et al.*, 2008, pp.27–35). It is from this point in time that the editors used this policy against their colleagues in order to argue for or against editorial changes. Since then, the act of attribution has seen a steady increase (Beschastnikh *et al.*, 2008).

As part of NPOV, it proclaims that original research, personal experiences, interpretations, and opinions do not belong in Wikipedia's articles. As Beschastnikh *et al.* (2008, p. 31) report, the policy of no original research was first written in September 2003 by Jimmy Wales: "If your viewpoint is held by an extremely small minority, then whether it's true or not, whether you can prove it or not, it doesn't belong in Wikipedia, except perhaps in some ancillary article. Wikipedia is not the place for original research." To this end, if a citation leads to original research that has not been published by a peer-reviewed or trusted source, then the content is illegitimated by not following the rules of NPOV. It is argued by Viégas, Wattenberg & McKeon (2007) that through these mechanisms, rules, and policies Wikipedia has successfully created articles of substantial quality.⁹⁸ Contrary to the claim that Wikipedia has removed authority from its pages, it can be stated that Wikipedians have created a system of knowledge representation that values the tradition of authorization and expert knowledge.

Alternatively, as opinion is ruled as inappropriate for Wikipedia's articles, the historian Roy Rosenzweig (2006) questions whether such an omission leads to quality articles on history. He comments that from a technical point of view, the kind of writing that history demands is difficult to do collaboratively (Rosenzweig, 2006, p.126). Not only does the process affect the style of writing, but the fact that Wikipedians mimic the "encyclopaedia voice" contributes to a standard and stolid writing style. While this has been deemed a valuable asset in encyclopaedias, Rosenzweig emphasizes that the lack of opinion and emotion is not necessarily beneficial for historical writing. Similarly, he remarks that "few would turn to Encarta or the Encyclopaedia Britannica for good writing" (Rosenzweig, 2006, p.132). That being said, despite the dry and emotionless style of the writing, he admits that "Wikipedia for the most part gets its facts right" (Rosenzweig, 2006, p.136) and is a reliable source of historical information.

Interestingly, he argues that history articles may in fact be of greater value to the contributors

⁹⁸Wikipedia has a special notification for articles that are exceptionally well researched and written. These articles are called *Featured Articles* and they are subjected to a rigorous process of evaluation and peer-review.

than their readers. He explains that participants in the editing process “often learn a more complex lesson about history writing — namely that the ‘facts’ of the past and the way those facts are arranged and reported are often highly contested” (Rosenzweig, 2006, p.138). The writing process itself becomes the most important aspect in this case, as it becomes a site in which intellectual conversation can lead to new understandings among its members. In fact, Pentzhold (2009) argues that Wikipedia goes beyond being “a platform to constitute and store knowledge, but a place where memory — understood as a particular discursive construction — is shaped” (Pentzhold, 2009, p.264). Instead of separating the discourse in creating an article, and the writing of the article itself, Wikipedia is a special kind of collective remembering where “the two modes do not displace each other. Rather, they are entangled. The stable article is generated parallel to the ongoing negotiations, which can furthermore retroact on already crystallized elements” (Pentzhold, 2009, p.264). This interaction is unique because it is articulated through “the wealth of arguments advanced during” the discussion of the article that are then reduced “to the set of validated statements that form the article” (Pentzhold, 2009, p.267). In other words, Wikipedia is not just an encyclopaedia but also a discourse between collective memories.⁹⁹

To summarize, Wikipedia is structured through a variety of organizational systems. At the software level, it is best described as a hypertextual database. This sort of connection facilitates the users’ ability to search the encyclopaedia. While this is the immediate link structure, editors have also taken it upon themselves to also connect articles through the creation of classifications. These larger subjects help the user navigate subjects of which they have no knowledge. While these systems describe the actual organization of the content, Wikipedia is also coordinated by the policies and mechanisms that Wikipedians agree to abide by.

⁹⁹ Another aspect of Wikipedia that has been questioned by scholars is that it is primarily a collection of linearly displayed words. There have been a number of studies that have created graphical representations of links between articles, concepts, and the histories of edit activity in order to achieve new understandings of how the content is created, related, and disputed. Viégas, Wattenberg & Kushal (2004) created a visualization of the edit history of an article in order to see conflicts and content movements over time. Milne (2010) develops and tests a visual navigator for Wikipedia that relies on the concepts of distance and scale to convey levels of article connectedness.

9.5.2 Connected Literature

The Wikipedia project states that it takes elements from “general and specialized encyclopaedias,¹⁰⁰ almanacs,¹⁰¹ and gazetteers¹⁰²”. As such, it differentiates itself from dictionaries, newspapers, and collections of source documents, as the functions of those works are covered by its Wikimedia sister projects (Wikipedia, 2005a). Beyond these empirically supported genres, Wikipedia is also metatextually connected to the *Encyclopaedia Britannica*. Comparative studies about Wikipedia most often use the *Britannica* as a barometer for change. For example, the often cited article “Internet encyclopaedias go head to head” (Giles, 2005) in *Nature* was the first to describe Wikipedia as near-parity with the quality of articles in the *Britannica*. Also Glott, Schmidt & Ghosh (2010) discover that Wikipedia’s articles are considered by readers to be in many regards higher than *Britannica*’s. Benkler stated at the time of his writing that Wikipedia “may well come to be seen as an adequate substitute for Britannica” as well as the Columbia, Grolier, and Encarta (2006, p.123). While Wikipedia is now the 6th most popular encyclopaedia site on the Internet, the *Britannica* (7312th) and Grolier (211,965th) are ranked substantially lower; the *Columbia* has been absorbed into Yahoo!; Encarta has been discontinued since 2009. Laudably, Benkler’s prediction has certainly come true. In addition to the usual suspects, Flanagan & Metzger (2011) studied the perceived credibility between Wikipedia, *Citizendium*, and the online *Encyclopaedia Britannica*. Their results showed that while Wikipedia is commonly used, it does not tend to be perceived as trustworthy or contain the same amount of expertise as its expert-generated competitors. In turn, when the source of the information was removed, children trusted Wikipedia articles significantly more. While these type of studies are important, Rut Jesus reminds us that regardless of trustworthiness and accuracy “Wikipedia is a top 10 website, and is widely used, cited or not” (2010, p.1). Beyond the *Britannica*, other encyclopaedias like the Online Columbia, Everything2, have been used in comparison to Wikipedia (Emigh

¹⁰⁰“An encyclopaedia [is] a compendium holding a summary of information from either all branches of knowledge or a particular branch of knowledge. Encyclopaedias are divided into articles or entries, which are usually accessed alphabetically by article name [...] encyclopaedia articles focus on factual information to cover the thing or concept for which the article name stands” (Wikipedia, 2001a).

¹⁰¹“An almanac [contains] tabular information in a particular field [...] arranged according to the calendar etc. Astronomical data and various statistics are also found in almanacs, [...] lists of all types, timelines, and more” (Wikipedia, 2003).

¹⁰²“A gazetteer [...] typically contains information concerning the geographical makeup of a country, region, or continent as well as the social statistics and physical features” (Wikipedia, 2001b)

& Herring, 2005). The overarching trend for connected literature is between Wikipedia and contemporary online encyclopaedias.

Alternatively, Kennedy (2009) suggests that Wikipedia is part of a “long and unbroken textual history”, one that has included the use of free information, collaboration and contribution, and blurring the roles of author and reader (p.229). She argues that many English-language texts build upon other texts through the “unifying tradition of textual borrowing and broad collaboration, resulting in a product that has rarely included originality among its central goals” (Kennedy, 2009, p.229). In addition, she notes that readers were often included in the writing process which was, “common in the eighteenth century, since the practice of selling a weekly or monthly subscription to a text printed in increments allowed the central encyclopaedist to solicit and incorporate contributions from the reading public” (Kennedy, 2009, p.236). Likewise, the idea of providing free information (as in free of copyright) is an old one. She explains that numerous “community-curated texts” have been “composed by legions of authors who receive[d] little acknowledgment for their work” (Kennedy, 2009, p.231). She makes the following point that “Chambers’ descriptions of encyclopaedic authorship, combined with his practice of incorporating submissions from readers, demonstrates that this sort of networked authorship did not spring into existence with the advent of the Internet. Chambers did not understand himself as a solitary author, nor did he function as one. Instead, he was deeply situated within a network of not only prior texts, but contemporaneous contributors” (Kennedy, 2009, p.239). Again she states: “Even a system as digitally dependent and apparently revolutionary as Wikipedia did not spring fully formed from the web. Instead, it draws on a long history of community-curated reference texts and the extensive prior development of the encyclopaedic form” (Kennedy, 2009, p.240). Thus Kennedy’s research supports the stance that the social and authorial mechanisms of Wikipedia are not new devices.

Another form of connecting literature on Wikipedia is the evolution of its citation practices. A graph in Beschastnikh *et al.* (2008, p.31) shows that half way through 2002 editors began to use the policy of citing one’s sources in the talk pages of articles. The page ‘Citing_sources’ in fact was first created in April of 2002. After an initial fervour, this practice declined until mid-2003 where it would steadily rise to become the most cited policy. Understandably, Wikipedians have put a lot of energy into ensuring that articles are cited properly. In Wikipedia’s case,

‘properly’ means to follow the rules stated on ‘Wikipedia:Attribution’, which explains that it is dedicated to ensuring that all “material in Wikipedia must be attributable to a reliable, published source” (Wikipedia, 2006).¹⁰³

9.5.3 Intended Audiences

One fact that cannot be denied is that Wikipedia is exceptionally popular. In fact, Wikipedia is the sixth most trafficked site on the Internet (Alexa.com, 2012a).¹⁰⁴ Part of its growth can be attributed to the “Google effect”. Explained by co-founder of Wikipedia, Larry Sanger who explains that “each time Google spidered [Wikipedia], more pages would be indexed; the greater the number of pages indexed, the more people arrived at the project; the more people involved in the project, the more pages there were to index” (Sanger, 2005). Being one of the most popular search engines on the Internet, Google has most certainly helped to establish and maintain Wikipedia’s popularity. What is important in Sanger’s quotation is that he alludes to the fact that merely having more people “arrive” at the site directly translated into more people being involved. In other words, those who are readers are expected to turn into authors of Wikipedia. This intention is reiterated by the tagline “the online encyclopaedia that anyone can edit”. While the English Wikipedia is the most popular version of the site, there are a total of 280 other languages that the project caters to in order to hold as true as possible to the tagline.¹⁰⁵ As Pentzold (2009) observes: “Instead of being a national-territorial project, it is divided into separate language editions. Hence, access depends not on nationality but on possession of the

¹⁰³For example, in the article *United.States*, the earliest form of attribution was a single link to a census PDF file in 2004. Then in 2005, a chart of census data is displayed and is captioned by a link to the US Census Bureau web site. Finally in 2006, the ‘United.States’ article contains an explosion of sources being cited as endnotes under the heading ‘References’. At first, these references are displayed in a single column but later in 2006 are displayed in two columns. Since 2006, this has been the standard format for references in the article ‘United.States’. However, each article has a different speed for adopting such policies as well as favoured sources of authorization. As of yet, the author of the current thesis has not found any studies that examine policy adoption rates per article, or which kinds of sources are heavily relied upon.

¹⁰⁴In an interview with *Wired*, Jimmy Wales claims that Wikipedia is in fact the fifth most popular web site (Solon, 2011).

¹⁰⁵The main languages that contain 1,000 000+ articles are English, German, and French. Article counts of 100,0001,000,000 include 37 languages. Eighty-four languages have between 10,000 and 100,000 articles, the remaining 175 languages have less than 10,000 articles each. (http://meta.wikimedia.org/w/index.php?title=List_of_Wikipedias&oldid=30291660October29, 2011) This is an unprecedented number of languages that one encyclopaedia caters to. Still, while there is a large scale effort to create interwiki translations of content, most articles differ greatly between languages. For instance, not only is content different between languages but Pfeil, Zaphiris & C.S.A. (2006) observe in their study that there are significant cultural differences in the editorial activities of different linguistic populations.

requisite language skills” (p.265).¹⁰⁶

Wikipedia is also inclusive given that anyone with electricity, a computer, an Internet connection and a browser can access and edit its content. Be that as it may, internetworldstats.com reports that “only a sixth of the world’s population is connected to the internet and that proportion is massively skewed, with Europe, North America and Asia making up 80% of those users” (Carnall, 2009, p.38). While computers are near ubiquitous in many countries, there are many that are still not connected. In addition to these statistics one American study observed that 36% of Americans online consulted Wikipedia and of those who do, a large portion of them are “well-educated and current college-age students” (Rainie & Tancer, 2007).¹⁰⁷ These numbers indicate that there are still some technical and social barriers to overcome if Wikipedia wants to uphold its claim that it is an encyclopaedia for anyone.

However, in terms of traditional concepts of an encyclopaedic audience, Wikipedia emulates the *Britannica*’s stance on the matter. As noted by the ‘Summary style’ article, Wikipedia states that it “is not divided into a macropaedia, micropaedia, and concise version, as is the Encyclopaedia Britannica” and as such it “must serve all three user types in the same encyclopedia” (Wikipedia, 2004a). The three types of users being spoken about are the three kinds that Hutchins describes in the 1984 edition of the *Britannica* (Hutchins, 1984, p.xi). This rationale for Wikipedia is thus used to explain why Wikipedia uses summaries to attend to different levels of detail as required by its users.

Another aspect of Wikipedia’s audience is the often discussed blurring of roles between readers and authors. This blurring is predicted by a 1989 article by Smith and reiterated by Stvilia, Twidale, Gasser & Smith (2005) when they remind their readers that not only would future encyclopaedias embrace hypertext but that “author and reader roles [would] be blurred and author contributions [would] be augmented by reader annotations” (1989/2005, p.2). As the discussion of Wikipedia has shown, this prediction has certainly been true. One such

¹⁰⁶Pfeil, Zaphiris & C.S.A. (2006) study the cultural differences between the German, French, Dutch, and Japanese Wikipedia versions of the article ‘game’.

¹⁰⁷Dooley (2010) argues that usage by students has created a pedagogical conflict within universities. Curiously, while there are official anti-Wikipedia policies at some universities, Dooley’s study concludes that many professors use Wikipedia as a source for information as well as a teaching aid. In support of academics being an audience of Wikipedia, Brazzeal (2011) analyzes the instances of Wikipedia within chemistry journals and finds that although it is not often used as a source, when it is, it is cited as credible. So as with previous encyclopaedias, the content appears to be especially useful to both students and academics alike.

consequence is free-riding.¹⁰⁸ Rahman (2006) contends that while many contributors work tirelessly to create a valuable encyclopaedia, the quality of Wikipedia articles is lowered by the substantial amount of free-riding. Antin & Cheshire (2010) disagree with this assessment. They argue that in many cases, some users of Wikipedia are unaware of their ability to contribute. As the authors note, it is often assumed that readers who do not contribute are free-riding. Nonetheless, the act of reading is recorded as a visit for articles, indicating the popularity of an article. Thus, the authors suggest that in many cases reading is a measurable act of peripheral participation, one that influences the actions of Wikipedia's editors.

Another form of activity that the audience participates in is monetary contribution. Even though it is free, the entire Wikimedia Foundation costs \$18,000,000 USD (Wikimedia Foundation, 2011) to operate with the majority of its capital resources coming from contributions, making up \$23,000,000 USD.¹⁰⁹ Despite the millions of dollars in costs to keep Wikipedia operating, the entire operation costs one third less than it did to just print the Fourteenth Edition of the *Encyclopaedia Britannica*.¹¹⁰ Obviously, the cost of Wikipedia's content costs nothing for the organization since it is completely voluntary.

9.5.4 Authors

Benkler lauds that what Wikipedia offers, unlike its expert-based competitors, is the opportunity for individuals to participate in the "writing" of their culture and to emphasize topics that are important to them (2006, p.293). While Wales holds a pseudo-oligarchy¹¹¹ that can ban disruptive users, the majority of coordination, as Benkler observes "relies instead on social norms to secure the dedication of project participants to objective writing" (2006, p.72). With the exception of a handful of cases dealing with living biographies,¹¹² coordination of the English

¹⁰⁸Free-riding is the act of taking advantage of the efforts of others.

¹⁰⁹The Financial Report for the Wikimedia Foundation for the year ending June 30, 2011 does not divide the costs up by project. Thus, it can be assumed that the Wikipedia projects cost less than the reported eighteen million dollars.

¹¹⁰Consider that the Fourteenth Edition of the *Britannica* cost \$2,500,000 USD to publish in 1929 and the "low price" of \$67.50 USD to buy the set during the Great Depression. Taking inflation into consideration, that would amount to \$33,170,906 USD and \$895 USD, respectively in today's dollars (Kogan, 1958, p.219 & p.237). Inflation calculation provided by United States Department of Labor, Bureau of Labor Statistics: http://www.bls.gov/data/inflation_calculator.htm

¹¹¹Wales describes his position as one that is closer to that of the British monarchy and that "my power should be (and is) limited, and should fade over time" (Reagle, 2007, p.150).

¹¹²Myers' article "Wikimmunity" (2006) discusses the legal implications of defamation on Wikipedia and how it has responded to these issues.

site is entirely left up to its fifteen million contributors (Wikipedia, 2012b). How this occurs has been heavily studied by researchers.

What Benkler finds surprising is that the social norms all come from “a large and geographically dispersed group of otherwise unrelated participants” (2006, p.74).¹¹³ Given the context that Wikipedia has the goal of representing the whole of knowledge, it is perhaps not as surprising as Benkler believes that the diversity of participants come together to share their knowledge.

Another aspect of Wikipedians that is often analyzed is the explanation of why people volunteer their time and energy. In a self-reported survey at the WikiMania’09 conference Wikipedians were asked why they contribute. The results showed that 72 per cent of them volunteered because they “like the idea of sharing knowledge and want to contribute to it” (Glott, Schmidt & Ghosh, 2010, p.9)¹¹⁴ This sentiment is summarized by Lévy: “No one knows everything, everyone knows something, all knowledge resides in humanity” (Lévy, 1994, p.13). There is perhaps not a better description that summarizes the ethos of Wikipedia.

What is curious about Wikipedians in terms of encyclopaedism is that there appears to be more energy invested in organizing people and activities than in actually writing articles. The often cited article by Kittur, Suh, Pendleton & Chi (2007) finds that as discussion pages are used over time, they have continually evolved to meet the editor’s needs for collaboration. In fact, while the amount of article-based activity has been decreasing, there has been a marked increase in activity on policy, rules, and talk pages. The authors conclude that this shift in activity indicates that the majority of contributors’ energy is now directed at self-organization.

It follows then that Wikipedia meshes the encyclopaedic voice of the article with that of the conversations of opinionated individuals. Although the project pages are written with the stoic and NPOV encyclopedic voice, the very presence of such content has been argued and negotiated

¹¹³The term “wikinomics,” coined by Tapscott, is used to explain “the process by which conscious social actors of multiple origins bring to others their resources and beliefs, expecting in return to receive the same” which would lead to Tapscott’s altruistic end of “the ancestral fear of the other” (2006, p.40). Lévy also describes that the Web can alleviate this fear of the other: “Who is the other? Someone who has knowledge. And moreover someone who knows what I do not. The other ceases to be the terrifying, threatening figure he is now; like me, he is ignorant of much and possesses limited knowledge” (1994, p.12).

¹¹⁴Conversely, 69 per cent of the respondents of the same WikiMania’09 survey responded with “I saw an error I wanted to fix”. Clay Shirky (2010) has written extensively on the subject and coined the term *cognitive surplus* to explain how some people are in social and economic positions where they can afford the time and energy to be part of this kind of community.

using the space dedicated to opinion. However, while they serve as a space of talk, they are for the most part, an internal function used by Wikipedians, and are not an encyclopaedic organization of arguments for and against the existence of controversial facts and positions.

Bazerman explains how there is a difference between these esoteric discussions of a community and the texts that it produces. Like encyclopaedias, Bazerman suggests that the communities that create scientific articles are predominately concerned with making texts that contain statements of knowledge. While these texts exist in physical forms, the social actions of the community are found in the negotiation of what statements are considered valuable (1988, p.22). He argues that once this “esoteric knowledge” is accepted by the community, it can then be used outside of the community in the form of a published text, one that is “held accountable and becomes the reference point for future discussion” (Bazerman, 1988, p.22). Thus in many ways the discussion and history pages serve as a backchannel to the creation of the encyclopaedia.

Perhaps the most revealing article that analyzes talk amongst Wikipedians comes from Jesus, Schwartz & Lehmann (2009). This study analyzes Wikipedian discussion pages for characteristic networks, or clusters, of users. The researchers observe and identify three kinds of user clusters; isolated, shared-interest, and non-content bounded clusters. Isolated clusters are observed in articles on Mormonism, Scientology, and Zionism. The researchers note the coincidence that these groups are known for their isolationist tendencies. Shared-interest clusters are also observed as a group of users that are consistent among a number of articles. The third group, non-content bounded clusters, tend to consist of the same users editing a diverse collection of articles. The explanation the authors offer is that these users are tied together by their shared activities and not a common interest. The benefit of this study comes from the first steps in being able to identify group activity types by looking at the discursive connections between users.

Implicit in being part of a community is also taking up roles that are attached to specific activities. As the Wikipedia page ‘Special:Statistics’ records, there are a great number of kinds of users.¹¹⁵ Of these users, the unregistered or anonymous users are often given special attention from researchers and critics alike. Both Kittur, Chi, Pendleton, Suh & Mytkowicz (2007) and

¹¹⁵Wikipedian user types: account creators, administrators, blocked users, bots, bureaucrats, checkusers, confirmed users, edit filter managers, file movers, founder, importers, ombudsman, oversighters, researchers, reviewers, rollbackers, stewards, transwiki importers, and unregistered users.

Nemoto, Gloor & Laubacher (2011) find that the the highest quality of edits were enacted by two groups “registered editors who make many edits and anonymous editors who make few” (Nemoto, 2011, p.2). Interestingly, Kittur, Chi, Pendleton, Suh & Mytkowicz categorizes these two groups as “elite” and “common” editors, alluding to the Marxist notion of the bourgeois. A similar study was conducted by Laniado & Tasso (2011) who seek to find out who makes the most substantial and long lasting writing on Wikipedia. In posing their question in this manner, they describe that the social editing of Wikipedia “resembles the collaborative writing of a scientific paper” (Laniado & Tasso, 2011, p.201), one where there is a definitive author to a text (or at least a hierarchy of writers). They use the same terminology as Kittur, Chi, Pendleton, Suh & Mytkowicz (2007) in that there is an “elite” core of users and the usual users who are the “bourgeoise”. They take this position even further by stating that any user who merely edits a page is not considered a proper author. In Laniado & Tasso (2011), the binary is reiterated when they describe in their methodology that they discarded data dealing with anonymous users (p.178). Even though anonymous users make up substantial material and ideological aspects of Wikipedia, these researchers have chosen to ignore them or at best, reduce their importance to the project. By looking at Wikipedia in this manner, these researchers are reinforcing the problematic belief in a “capital A author”, an author that is original and powerful (Kennedy, 2010, p.303). Kennedy explains that applying the concept of “Author” to Wikipedia is misleading. She argues that authors of Wikipedia exist within a spectrum of agency and not a simple binary. There are different roles, tasks, and even personhoods that authors are a part of on Wikipedia. Reminiscent of Latours’ actors, Kennedy describes how even bots throw the entire concept of authorship into contention.¹¹⁶ However, these activities were once relegated to editors and copy editors, so in a way, their activities are not new, only the actors. Therefore, asking the question of “[w]ho are the *real* authors of Wikipedia?”, as Laniado & Tasso do (2011, p.201), invalidates the variety of activities that the community participates in.

Rut Jesus (2010) offers an alternative scheme to categorizing contributors. Following his article on bi-partie collaborative networks, Jesus discusses the cognitive differences between the substantial editors and those who make only minor changes. Jesus correlates these two groups

¹¹⁶Bots perform the mundane but necessary actions of such an enormous corpus (correct spelling, finding dead links, linking articles, etc.).

as working with varying degrees of cognition for planning and cognition for improvisation. He prescribes that Wikipedia's success comes from its modularity.¹¹⁷ From these terms of planning and improvisation, Jesus promotes that minor edits (regardless of membership or anonymity) should be considered contributions, cognitive artifacts, but their authors do not warrant the status of contributors. While the previous studies discount anonymous users for sake of convenience, Jesus provides a logical division between user types.

Studies that are concerned with social roles often look for forms of leadership that define the social milieu. Castells points out that 'the community accepts the hierarchy of excellence and seniority only as long as this authority is exercised for the well-being of the community as a whole' (2001, p.48). This is indeed what Thom-Santelli, Cosley & Gay (2009) explore in their article. They find that despite the goal of a shared enterprise, many contributors develop a sense of "territoriality" over the articles they work on. The results of this are double-edged. On one hand, the "lead" editor removes vandalism, maintains the quality of the article, and encourages participation. On the other hand, the authors observed that defensive positions may be established that deter group collaboration. The authors conclude that despite these issues, territoriality remained a necessary practice for a hierarchically collaborative space like Wikipedia.¹¹⁸

While the previous article argues for the necessity of hierarchal social structures, Zhu, Kraut, Wang & Kittur (2011) find that such structures are always shifting. Their results show that social leaders are not necessarily just administrators, but exist at all levels of membership. They state that since there is not a set hierarchy there are shifting responsibilities (like supervising, rewarding, warning and socializing workers) between leaders. These activities have also been studied as newcomers become Wikipedians. Bryant, Forte & Bruckman (2005) describe that through the practices of information gathering, policy adherence, and technological mechanisms like watchlists "the structure and character" of the community is realized. The level of participation surrounding Wikipedia is a good example of Lévy's collective intelligence at work.

¹¹⁷From *The Wealth of Networks*: "Modularity" is a property of a project that describes the extent to which it can be broken down into smaller components, or modules, that can be independently produced before they are assembled into a whole. If modules are independent, individual contributors can choose what and when to contribute independently of each other. This maximizes their autonomy and flexibility to define the nature, extent, and timing of their participation in the project. (Benkler, 2006, p.100).

¹¹⁸Interesting to note, as Reagle (2007) discusses, is that while there are many policies and rules that govern Wikipedians, none of them tackle the issue of leadership (p.147).

Contrastingly, while many Wikipedians work toward a common goal, many use Wikipedia for alternative authorial activities. The presence of vandalism has often been used to discredit the quality of Wikipedia. As Kennedy (2009) notes, what has been considered acts of vandalism by the majority demonstrates “the ambiguous nature of agency by refusing to recognize or actively working against the symbolic constraints of form and community” (p.238). In fact, as Kennedy argues, vandalism represents a participatory audience that is creating alternative “points of articulation” (2009, p.238). However, even this activity has been subject to policing by the community.¹¹⁹

It is evident that leadership, policy, and surveillance raise the level of article quality.¹²⁰ Indeed, Flanagin & Metzger (2011) report that Wikipedia has become a common source of reference. Thus for millions, the question of quality has been settled and that “the interesting questions are elsewhere” (Shirky, 2008, p.117).

While Benkler, Shirky, and Castells are flag-wavers for the digitization of knowledge and projects like Wikipedia, other authors are not so favourable. Perhaps the straw man of this group of authors is computer scientist Jaron Lanier (2010). He has an overall fear that we will soon, if not already, be subsumed by the mob of anonymous amateurs. He attacks projects like Wikipedia and Linux for emphasizing the value of the crowd, alluding that such attitudes deemphasize individuals and thus creates an environment in which humans “revert to bad moblike behaviors” (Lanier, 2010, p.19). Similar sentiments, (albeit less polemic) are reiterated in Tapscott’s (2006) *Wikinomics* and Andrew Keen’s *Cult of the Amateur* (2007). Tapscott insists that “credentialed knowledge producers share the stage with “amateur” creators who are disrupting every activity they touch” (2006, p.11). In a similar vein, Keen cautions that the amateurs of Reddit, Youtube, and Wikipedia should not be trusted merely because of their status as amateurs (Keen, 2007, p.95).

¹¹⁹Priedhorsky *et al.* (2007) find that despite the amount of vandalism that is likely to occur on Wikipedia, it is often reverted in a matter of minutes. For example, they state “42 % of damage incidents are repaired within one estimated view” (Priedhorsky *et al.*, 2007, p.265).

¹²⁰The authors Stvilia, Twidale, Gasser & Smith (2005) attempt to determine the quality of Wikipedia articles. They adopt the scheme proposed by Crawford (2001) that identifies how comprehensive an encyclopaedia is by how well it solves questions of ready reference, general background information, and pre-research information “leading to more targeted and detailed sources” (Stvilia, Twidale, Gasser & Smith, p.2). They take this scheme in conjunction with evaluating an encyclopaedia’s scope, format, uniqueness, authority, accuracy, reliability, objectivity, currency, accessibility, relevance, and cost. With this study, the authors find that Wikipedia adds valuable data to understanding what people consider a “quality” article.

The assumptions that Tapscott, Keen, and Lanier make are that these sites, for the most part, are populated by amateurs. Given the rules of verifiability and “no original content” on Wikipedia, this argument appears moot since resilient edits are often linked to expert sources. Specifically, because anyone can contribute to Wikipedia, there must necessarily be some form of filter to insure the validity of content. Original research that is not published in a peer-reviewed or reputable source is considered as unverified knowledge. Therefore, while the writing may be made by “amateurs”, they are still supported by expert sources.

Furthermore, a self-reporting study of Wikipedia users places the percentage of experts to amateurs on articles at 77 per cent and 23 per cent respectively (Glott, Schmidt & Ghosh 2010).¹²¹ In math related subjects, the number of experts is as high as 90 per cent while the lowest number of experts, 67 per cent, was found in subjects related to people and the self. While the arguments of Lanier and Keen appear more reactive than reflective, there is evidence that important issues lie elsewhere.

What can be seen as one of the major obstacles to Wikipedia’s inclusivity is that it has retained the male-dominated authorship of previous encyclopaedias. For instance, in Glott, Schmidt & Ghosh’s demographic study (2010), the makeup of Wikipedians is overwhelmingly male (two times more than females), English-speaking, and from the ages 22 to 49. Likewise, there is a significant correlation between the age of these males and their preferred subjects. The study reports that males older than thirty tend to contribute to geography, history and cultural sciences; users younger than thirty tend to contribute to natural sciences, mathematics, logic and technology. With the exception of cultural sciences, these subjects have long been the subjects of encyclopaedic writing. A more recent study also concluded that the number of Wikipedian male contributors had only dropped to 82 per cent (Antin, 2011). While there is much to be said about the innovating and self-organizing networks of Wikipedians, one cannot help but be wary of repetitions of the colonial “old boys club” that continually threatens the ethos of an encyclopaedia written by anyone.

In fact, Flöck, Vrandecic & Simperl (2011) argue this point in their article that observes that dialogue on the Web often reinforces the homogeneous beliefs of a community. The authors note

¹²¹The authors of this self-reporting survey admit that explicit categories of education and profession would need to be validated in future studies to offer greater validity to these statistics.

that the inclusion of different opinions are often silenced by a majority consensus and content that has existed longer is harder to remove. The explanation for this is that this content has been conceded by many editors and readers and is therefore perceived as reputable information. These authors conclude that there exists substantial barriers that block new perspectives from being accepted in Wikipedia “even if they objectively contain useful information” (p.5). The findings of Flöck, Vrandecic & Simperl (2011) conflict with the claims of Benkler, Castells, and Lévy in terms of Wikipedia being a site of open discussion. While Wikipedians put a lot of energy into coordinating talk, the product that they create may be more homogenous than one might expect. Wikipedians themselves have noted how a substantial amount of content is dedicated to “geek priorities” (Rosenzweig, 2006, p.127), topics that are dominated by a relatively small portion of the general population. Beyond the esoteric, Wikipedia has also seen a huge influx of articles on popular culture, noting that users of Wikipedia are interested in topics that obviously reflect their own interests.¹²² These results are perfectly in keeping with Benkler’s claim that individuals are authoring and emphasizing topics that are important to them.

9.6 Synthesis of social contexts

The purpose of Section 2 of the analysis has been to reveal the epistemological choices of encyclopaedism by exploring how the genre has realized its communicative purposes through the four contexts of lexicon, connected literature, intended audience, and authors. As described by genre theory, this exploration has revealed that encyclopaedism has constantly evolved and changed its conventions to adapt to new situations within recurrent contexts. However, to understand how encyclopaedism as a network is in conflict with wikis on Wikipedia, the current study must attempt to define some of the discriminative strategies that have been consistent in realizing the genre’s goals.

The following subsections first provide an overview of the evolution of the genre through the four contexts. Following this overview, the analysis outlines five discriminating strategies (utility, systematic organization, consistency, authority, and trust in experts) that reveal the

¹²²According to the wikistats site, <http://stats.grok.se/en/201011/> for the month of December 2010, the most popular articles included WikiLeaks (at the time, a popular news item), Christmas (holiday), Facebook (social software), Julian Assange (personality behind WikiLeaks) Glee (TV series), Tron (film), Google (popular search engine), Justin Bieber (musician), and Nicki Minaj (musician).

epistemic values of encyclopaedism.

9.6.1 Overview of contexts

Lexicon: Some encyclopaedists (like Chambers) prided themselves in their alphabetically arranged work. Alternatively, others (like Coleridge, editor of the *Metropolitana*) held that topical arrangement, similar to the liberal arts, was the best way to arrange information. He also chided that alphabetically arranged encyclopaedias are a “‘huge unconnected miscellany’ of accidentally aligned articles” (Sullivan, 1990, p.320). This attitude is reiterated among rival encyclopaedias, even within various editions of the *Britannica*. However, Sullivan defends that even though the alphabet was an implicit aspect of organizing the new encyclopaedias, it was not responsible for the overarching order (1990, p.320). Understandably, these formal matters of organization were real concerns as the sheer size of encyclopaedias began to increase almost exponentially.¹²³ Through this period we see a constant debate between rival encyclopaedias and between subsequent editions which vie for the reading public’s attention. Ephraim Chambers claimed that his work was a condensation of all useful knowledge, and thus a suitable substitute for libraries (Yeo, 2007, p.47). The 1842 Preface of the *Britannica* explains that the encyclopaedias that dominated this era “were intended to comprise, within a moderate compass, a series of systematic digests applicable either to some of the great departments of knowledge and art, or to the whole body of learning” (EB, 1842, p.vi). Regardless of topical or alphabetical order, Sullivan argues it was the response to the geographic and economic expansion of European colonialism that spurred the need to “refashion the classification of human knowledge” (Sullivan, 1990, p.316). Sullivan indicates that the “culling and organizing” of encyclopaedias is directed by an “overarching goal”, one that makes the relationship between truth and knowledge meaningful (1990, p.317). Despite the various attempts at providing the most useful means of organization and a renewed vigour for unifying knowledge, Chambers, Diderot, and d’Alembert each “confessed that there was no single way of classifying knowledge” explaining that “since humans made knowledge (as opposed to the objects and structures of the natural world, created by God) we could classify it in ways that suited our purposes” (Yeo, 2007, p.55).

¹²³In the case of the *Encyclopaedia Britannica*, the First Edition only had three volumes while the Third Edition, only twenty years later, had eighteen.

As such, traditional views of encyclopaedic organization have tended to revolve around the alphabetization and topical systems of organization. However, the current study has described several organizational schemes that have been crucial in providing cognitive and material connections between knowledge. For example philosophical (Bacon) and topical (disciplines), alphabetical (A-Z), hypertextual (cross-references), intertextual (citations), and database systems (wiki) have each been used to distribute, organize and provide access to content. As such, in all four encyclopaedias, there is not one primary form of organization but an assemblage of organizational systems. For instance, the *Cyclopaedia* utilizes alphabetical, hypertextual, intertextual and a measure of philosophy to organize this encyclopaedia. The *Encyclopédie* uses alphabetical, hypertextual, intertextual as well as explicit philosophical organization. The *Britannica* has from time to time used a philosophical organization but more often relies on alphabetical, hypertextual, intertextual, and topical organizations. Recently, both the *Britannica* and Wikipedia fold these three organizational types into a database organization which is organized alphabetically, connected hypertextually, and is accessed through keywords (topic). As these differences display, the tactics that each encyclopaedia uses to organize its content has differed in both scope and intensity. Therefore no one system can be considered more encyclopaedic than any other.

From the beginning, encyclopaedias have relied on summary to convey knowledge. However these summaries have been augmented from time to time by treatises. In these early examples, it was quite common to find opinion and argumentation about subjects.¹²⁴ This makes sense since the *Britannica* was often sold on the premise that its contributors were intellectual celebrities and whose opinions were valued by both the encyclopaedia and the public. Paradoxically, in the same editions that laud the opinions of intellectual celebrities as a selling point, the rest of the encyclopaedia was deemed to relate encyclopaedic knowledge as objectively and as unbiased as possible. Baynes states that the duty of the *Britannica* “is to give an accurate account of the facts and an impartial summary of results in every department of inquiry and research” (1878, p.vii). This convention continued when Preece described that information from the *Britannica*

¹²⁴Commenting on the Ninth edition the Edinburgh Review states that “There never was a time when the results of science and the results of history have been more ably presented and preserved, although, as is inevitable, it exhibits and provokes diversities of opinion on many subjects” (Kogan, 1958, p.61), illustrating that a neutral point of view was not always a prerequisite to encyclopedic writing.

should be provided to its readers in a way that was integrated, coherent, limiting the amount of duplication of information, written objectively and with neutrality in mind (1967, p.xv). In Wikipedia the pursuit of neutrality has been further embedded in the culture as an explicit core policy. As a result, the appeal to neutral and objective summaries has long been identified as the encyclopaedic voice.

Connected literature: The connected literature context illustrates how the four encyclopaedias have tightened the bonds between each other and other texts through metatextuality and intertextuality. In the first instance each encyclopaedia compares itself to, comments on, or aligns itself with previous encyclopaedias. Chambers aligned himself with famous lexicons and dictionaries; the *Encyclopédie* was itself originally a translation of the *Cyclopaedia*; and numerous prefaces of the *Britannica* lambasted and praised those two predecessors; and Wikipedia participates in being connected to these encyclopaedias through academic research.¹²⁵ In all, these self-endorsed and scholarly connections have tightened the generic links between the encyclopaedias and have justifiably instilled a canon that any new encyclopaedia will thus be compared to, compete with, and incited to improve upon.

The intertextual ties used in each encyclopaedia explain how the practice of citation is used to establish authority and verifiable knowledge. While the *Encyclopédie*, the *Britannica*, and Wikipedia consistently use citations, the *Cyclopaedia* was less rigorous with this practice. As expressed by Chambers, he did not need to always cite the works he borrowed and summarized because they were in the service of the public. This view of citation was rejected by Diderot & d’Alembert when they decided that the practice of citation was as important to connecting knowledge as were Chambers’ cross-references. Thus the *Encyclopédie* not only affirmed “the interrelation of knowledge” but also “account[ed] for its origin” (Ferris, 2003, p.1260). This approach to citation has continued steadily and has formed the foundation for establishing the verifiability of knowledge on Wikipedia.

Intended Audiences: Throughout the evolution of the genre, intended audiences have played a pivotal role in determining the utility of encyclopaedias. In the beginning, encyclopaedias

¹²⁵The current thesis itself is an example of such academic activity that places Wikipedia squarely beside its predecessors.

differed themselves from summas by conceiving their audience as not only the clergy, but also doctors, lawyers, and women (Burke, 2000, p.189). Likewise, Yeo comments that encyclopaedias differed from previous genres because they “were commercial ventures pitched to a readership wider than the classically educated elites of church, state, and university” (2007, p.50). Indeed, a fundamental difference between modern encyclopaedias and the summas that preceded them came from the fact that readers were primarily part of an economic model and not spiritual fidelity.

One of the most prominent intended audiences has been the expert or specialist. All of the encyclopaedias studied up to the 1973 edition of the *Britannica* spoke of the fact that specialists would use encyclopaedias to keep abreast of the latest developments in their field. This was supported by the fact that new and original scientific research was written about in the nineteenth century editions of the *Britannica*. However, catering to specialist audiences was retracted by 1973. This edition of the *Britannica* stated that since it was only a report on confirmed knowledge, new knowledge would rarely be added for the benefit of a specialist community.

In opposition to expert audiences are the non-specialists, the “average man” “the layman”, or the “man of ordinary parts”. In constructing the audience this way, encyclopaedias envision taking the esoteric knowledge of experts and representing it to the public at large. This was of course useful as it was on this platform that the *Britannica* established itself as a commercial success, making itself a common sight in the living rooms of the general public. In this respect the audience was not only considered to be readers of the encyclopaedia but also consumers and subscribers. In the early encyclopaedias subscribers were so valued by the editors that each subscriber was listed within the encyclopaedia itself.

While experts and average men make up the majority of the intended audiences of these encyclopaedias, there has been a shift in intentions since the 1984 edition of the *Britannica*. After this point it has been more common to describe the audience as users. This trend has followed through with Wikipedia but with a crucial difference. A user of Wikipedia spans the spectrum of textual agency, from the casual reader to the prolific writer. As such, Wikipedia now considers its audience to include not just readers but also the editors and contributors of encyclopaedic content.

Common to much of these perceptions of audiences has been that the majority of users are students and men. These two audiences have consistently been described as the primary audiences. Contrastingly, the first mention of women as an intended audience was not made explicit until the 1984 edition of the *Britannica*.

What these various roles of audiences reveal are the specific relationships that producers of encyclopaedias envision. For the most part, encyclopaedias have catered to students, specialists, and average men. These identities have been used to validate the educational purposes of the encyclopaedia and to position encyclopaedias within the scientific community. In most cases, this has been described as a male-oriented genre. As time has passed, different audiences have been added to this mix for economic reasons (subscribers), social reasons (women), and even production reasons (contributors). Interestingly enough, the Wikisym '09 Survey (Glott, Schmidt & Ghosh, 2010) of Wikipedians reports that men, students, and experts continue to be three groups that dominate the creation of Wikipedia.

Authors: As the review of contexts has shown, there has been no clear convention as to whom should write encyclopaedias. The writing of the first encyclopaedia was taken on by a single man, Ephraim Chambers, who was more responsible for summarizing previous texts than for writing a new one. Later, Diderot & d'Alembert enlisted 140 contributors to write wholly new texts on new subjects. The use of contributors in this case included both experts and intellectual celebrities to draw in more subscribers. In the *Britannica*, there was a period where new original scientific research had a proper place between its covers. As such, authorship was relegated to experimental researchers in burgeoning fields of science. However, the inclusion of new science was later removed as the *Britannica* moved toward being a report on the state of knowledge rather than being a space to publish recent discoveries. By the turn of the twentieth century, the reach of authorship had also begun to shift from the exclusivity of men to the inclusion of women (although in relatively low numbers). By the mid-point of the century, it was necessary that authorship was also shared internationally. Wikipedia has embraced and extended some of these qualities of authorship as displayed by both its international membership and 231 language-based Wikipedias. However, despite this quest to include more contributors, male Wikipedians continue to unjustifiably outweigh female Wikipedians.

The evolutions of authors has also happened quantitatively. In the past three centuries the number of authors involved has continually risen. From one in the *Cyclopedia*, to 140 in the *Encyclopédie*, to 5000 in the *Britannica*, and 18 million registered contributors on Wikipedia (with an unknown number of unregistered contributors). In view of these evolutions of encyclopaedic authorship, the tactic of involving more people can be explained by the goal of including more knowledge, which is directly related to the communicative purpose of the genre.

Another aspect of encyclopaedic authors is that contributors tend to participate more for altruistic reasons than economic compensation. While Chambers was the sole editor of his encyclopaedia, Diderot & d'Alembert appealed to numerous volunteers to contribute. During the reign of the *Britannica*, contributors were more likely to be paid than to be volunteers. However, their compensation was often meagre, playing to contributors' altruistic belief in the encyclopaedic enterprise. Of course, Wikipedia continues this trend of altruistic volunteers as they are wholly responsible for the production of the content on the encyclopaedia.

Of the discussions surrounding encyclopaedias, there is the misleading commentary that assumes the authors of encyclopaedias have always been experts and that Wikipedia is unique in that it undermines this authority by allowing anyone to contribute. This is not entirely true. At the beginning of the genre we saw that the editors were responsible for summarizing and writing the encyclopaedia and were not in themselves experts. In this sense, Wikipedia does more to carry the tradition than to upset it. However, what is interesting about Wikipedia is that while it devalues the role of expertise of its contributors, it has substantially increased the co-presence of experts through citation. In this way, the statement that Wikipedia undermines the authority of experts can only be true if authority is only confined to contributors. As has been explained by Sanger (2009), in order for Wikipedia to be regarded as a legitimate source of encyclopaedic knowledge, it has to retain the authority of experts even though its contributors are not necessarily experts themselves. Because of the practice of citation, only knowledge that can be verified through trusted sources has a chance of remaining on an article's project page. Because of these facts, Wikipedia has not escaped the authority of the expert but has conversely placed explicit trust in them.

What can be gleaned from this overview of contexts is that encyclopaedism has constantly evolved through the ways it organizes its content, increases of citation practices, the identities

that audiences are given, and the roles that its authors play. Within these evolutions there are a number of common and discriminative strategies that strive to achieve the communicative goals of the genre. As such, the following subsection outlines these strategies in order to reveal the epistemology of encyclopaedism.

9.6.2 Discriminative strategies and epistemic values

In the previous overview, the four contexts have indicated how encyclopaedism has evolved over time. However, in order for encyclopaedism to be considered a genre, it must illustrate that there is a consistency of discriminative strategies for achieving the communicative purposes. As such, the following subsections not only outline these consistencies but additionally highlight the epistemic rationales that guide the production of the genre. Therefore, the purpose of this analysis to reveal the epistemic values of encyclopaedism is made explicit in the following list.

Utility: The value of an encyclopaedia, as noted by several encyclopaedists, can be determined by the degree of utility it offers. As the opening sentence of the 1771 *Encyclopaedia Britannica* Preface states: “Utility ought to be the principal intention of every publication. Wherever this intention does not plainly appear, neither the books nor their authors have the smallest claim to the approbation of mankind” (EB, 1771b, p.v). Two hundred years later, the editor of the 1984 edition admits the same sentiment that “utility is in fact the name of the unit in which encyclopaedias must inevitably be measured” (EB, 1984, p.xi). Utility in this sense can be described using Jeremy Bentham’s famous maxim that utility is “the greatest happiness of the greatest number that is the measure of right and wrong” (Bentham, 1776, p.ii). In other words, Bentham explains that the principle of utility is the “property in any object, whereby it tends to produce benefit, advantage, pleasure, good, or happiness” (1823, p.3). In the case of an encyclopaedia, the happiness of the user is obtained by giving the least painful mode of access to the communicative purposes. As such, the utility of an encyclopaedia can be measured in two ways: by how it provides material access to the totality of knowledge; and how it provides cognitive access to understanding how that totality is connected. This view of encyclopaedism is maintained by Withers’ comment that accessibility and the utility that it grants became “the primary concern” of the eighteenth century encyclopaedias (Withers, 1996, p.279). Truly, this

principle can be seen at work in a number of transformations within encyclopaedism. For instance the early move from writing in Latin and Greek to vernacular languages make it easier for more people to access encyclopaedic knowledge. More recently the move from print media to the Web has provided encyclopaedias an unprecedented amount of textual space, the ability to update in real-time, and has reduced the amount of time searching for a subject. Each of these tactics allows a more efficient access to the topics and articles a user wants to learn about. However, the communicative purposes are not solely achieved through utility but through a constellation of strategies.

Systematic organization: What is common to all four encyclopaedias is that each has been thoughtfully organized. This aspect of encyclopaedism has been often noted and the current examination shows no reason to disagree. The importance of this strategy to the communicative purposes is that it achieves both totalization and education of knowledge by providing both a cognitive and material system by which each piece of knowledge is connected. On one hand the organization must address the relationships between different disciplines so as to inscribe in the reader a sense that all knowledge is connected. On the other hand a material system of organization allows users to locate, navigate, and explore the articles of knowledge. How this occurs is left up to the individual encyclopaedia to decide. As such there has been a variety of philosophical, alphabetical, hypertextual, intertextual, topical and database tactics to achieve this strategy.

Consistency: Each of the encyclopedias studied use summary and categorization as a tactic to actualize the communicative purposes. In doing so they reveal that encyclopaedism values knowledge that is consistently reduced to essential and defining qualities. In the case of categorization within traditional encyclopaedias that followed Bacon's philosophy, knowledge is validated if it fits within the three faculties of the mind: reason, memory, and imagination. In a similar fashion, Wikipedia states that the "central goal of the category system" is to link all articles through the "essential" and "defining characteristics" of a topic, and each article of that topic "must be able to stand alone as a self-contained unit" (Wikipedia, 2004c). In this way, only articles that are consistent with these qualities will be accepted, as the purpose is to provide

a system that connects all knowledge together. Therefore any topic or article that cannot be connected through this system is effectively inappropriate to the encyclopaedia.

Similarly, summary is the tactic of taking a large body of knowledge and reducing it to core aspects that are consistent with the systematic organization. As a material device, it aids in providing utility to the totality of knowledge by making it digestible. Yeo explains that “summary gave the only chance of making the essential knowledge of a range of sciences accessible to all” (2003, p.65). Additionally, the appeal to consistent qualities in the beginning was supported by the medium of print as it was costly and unwieldy to publish everything there was about a subject. As such, encyclopaedias used intertextual connections to open the knowledge up from its folded and reduced form. For Chambers, the only way to achieve his goal was to consistently reduce knowledge “to essentials” and collect it “in an accessible work” (Yeo, 2003, p.71). In the case of Wikipedia, summary is used more for the benefit of its users than a response to technological limitation. The article ‘Summary style’ explains that summarization is accomplished “by not overwhelming the reader with too much text up front, by summarizing main points and going into more details on particular points [...] in separate articles” (Wikipedia, 2004a). On a different level, the use of summary assumes that such a reduction is a legitimate form of representing the interconnections of knowledge. Philosophically then, encyclopaedic summary and categorization tactics are representative of the epistemic value of consistency. However this consistency does not stem from an explicit usage of a priori essences nor an adamant position that such categories are socially constructed. In each encyclopaedia, the choice to divide an article or to describe its core aspects have been deferred to the knowledge of experts. In the case of the *Encyclopédie*, Diderot & d’Alembert relied on the expertise of Francis Bacon. In the case of Wikipedia “a defining characteristic is one that reliable, secondary sources commonly and consistently define, in prose, the subject as having” (Wikipedia, 2006a). As such, the logic of encyclopaedism’s structure and textual form emerges from its trust in the authority of experts.

Authority: The creation of authority in the four encyclopaedias has been an important strategy to pursue the purpose of representing the whole of knowledge. Because encyclopaedias represent knowledge through summary, there must be some epistemic logic to why one detail of a topic

is more essential than another, and thus legitimates the need to be summarized. To do so, encyclopaedias have relied on the authority (not expertise) of contributors to make these decisions. The strategy of authority to validate the summary has been achieved in a variety of ways. It has been manifested through a spectrum of tactics, ranging from argumentation, opinion, and expertise, to objectivity and neutral points of view.

These tactics have been used at different times for different secondary reasons but all strive to address the totality of human knowledge. In some cases, like the treatises of the *Britannica*, this was achieved through opinion and argumentation, spelling out for the reader the logic of how one discipline is connected to another. In other cases it was through the use of expert contributors. The current trend on Wikipedia is to rely on the encyclopaedic voice or a neutral point of view. This authoritative tactic attempts to display a balanced view of all the different aspects of an article. Authority in this sense is the voice that totalizes, captures, and summarizes all bias, which is in keeping with the communicative purposes.

Trust in experts: In addition to legitimizing a particular summary, summaries must also be verified to be true so as to be representative of knowledge. To do so, encyclopaedias have put their trust in experts. Verification of truth has thus been attained by either summarizing a trusted source, by giving co-presence to a source through citation, or by giving an expert the role of contributor. While the issue of reliability of Wikipedia tends to revolve around the lack of authority given to expert contributors, this tactic is but one of several to verify the truth of a summary within an encyclopaedia.

9.6.3 Conclusions

Returning to Bazerman, he describes a genre as “a social construct that regularizes communication, interaction, and relations. Thus the formal features that are shared by the corpus of texts in a genre and by which we usually recognize a text’s inclusion in a genre, are the linguistic/symbolic solution to a problem in social interaction” (1988, p.62). In the description of how encyclopaedism emerged, the encyclopedists form a solution to problems of knowledge. They ask how can we find and use knowledge that is locked away in innumerable books? How can we connect related knowledge together so that one is able to find what one needs? How

do we incorporate new knowledge into existing knowledge? How is it best to organize this knowledge? How specific or general should an encyclopedia be? These questions and more demarcated the genre. Once a number of influential encyclopaedias established certain ways of solving these issues of social interaction, other reference works soon followed suit, stabilizing conventions while still providing new avenues to explore.

As such, the five discriminative strategies listed above can be viewed as constructing one particular solution to the problem of how to share knowledge. Each of these strategies set out the conditions in which to judge whether a statement is a valid contribution to encyclopaedic knowledge or not. As such they represent the epistemic values of encyclopaedism. Therefore, knowledge in this form must strive toward the two communicative goals of a total system of knowledge that is educational. Additionally, knowledge can only be considered encyclopaedic if it describes the essential qualities of subjects; is accessible through an established organization system; speaks from and with authority; and is validated by experts. Put together, these strategies make knowledge into a utility to be used.

What must be noted is that attributing these values to encyclopaedic knowledge is not entirely new as they have been, in large part, been described before.¹²⁶ However, the current study has uniquely refashioned the attributes into goals and strategies, providing a framework from which to identify how these values are part of a discourse that can methodologically be placed on the same level as the epistemic values of wikis.

10 Section 3: Textual structures

Rex Clark (2001) states that encyclopaedic projects are the manifestation of theories on knowledge. He asserts that they establish “what is knowledge, how it is organized, and what should be included or excluded” (p.462). In a similar fashion, Lévy imparts that it is important to understand the categories we use to classify knowledge because as “the world of categories becomes an object of cognition [...], the activity of knowing can therefore reflect itself” (2011,

¹²⁶In 1984, Hutchins proclaimed that encyclopaedias must cover the whole of knowledge; must speak “from and with authority”; must be summary in nature; be accessible through the structure of organization of content; be accurate; and be internationally composed (Hutchins, 1984, p.xi). With the exception of the first and last points, all of the encyclopaedias examined in this study have confidently met these requirements. Of course, while each encyclopaedia cannot be said to actually cover the whole of knowledge, as a communicative purpose, they certainly strive to do so.

p.80). Justifiably, the following paratextual analysis of the headings used in the *Britannica* and Wikipedia is useful as it investigates how such epistemological choices are activated at the textual level of discourse.

The following collection of data occurred between February 19 and June 22, 2011. In the case of data from Wikipedia, it was copied and pasted from HTML documents into a spreadsheet application. In the case of the *Britannica*, data from the years 1771, 1778, 1790, 1798, 1910 was collected from digital scans from online archives. Data from the years 1842, 1853, 1873, 1968, 1974, 1984, 1990 and 2005 were transcribed from physical books. Data from the the *Online Britannica* was copied and pasted from HTML documents. Once this data was either electronically copied or typed by hand into a spreadsheet application, it was manually coded by the researcher.¹²⁷ During the coding of the data, a number of inconsistencies were found.¹²⁸

¹²⁷In the first round of coding, it was not always apparent where some headings fit within the typographic hierarchy. For instance, a heading type was placed completely in the margins or shoulder of the text. In these cases, the term glossa was used as a placeholder for these kinds of heads. This term was used since the headings are paraphrases of the text. As Jackson explains in *Marginalia* (2001) “The gloss, in its primary sense, translates or explains foreign or obscure words; its expanded forms are the translation and the paraphrase. It operates at the most literal of levels, and aims to be faithful to the text it mediates” (p.45). After the initial coding, a thorough reading of the text was the only way possible for establishing where in the typographic hierarchy the glossa should appear.

¹²⁸*EB 1771–1798*: The publications during these years still used the typographic tradition of using the long s and the double s (ß) characters as well as c-t and s-t ligatures. These characters have been recoded into their modern English equivalents to retain spelling consistency between all editions.

EB London 1790: Art. 100 skips to Art. 111 in the source article.

EB Canada 1798: The data from this article is missing because the online source was inaccessible. It is assumed that the article does exist, given that the 1790 edition (which is the British version of the 1798 edition) is almost an exact copy.

EB 1878: The body and headings of articles vary in size by two points, but this change does not affect the logic of the hierarchy. It is assumed that such adjustment in size was an attempt to reduce the page length of the article and not a change in hierarchal order.

EB 1910: A change in body size (and subsequent peritextual elements) does not affect hierarchy. The choice for this change appears to be a choice in conserving physical page space, and not an attempt to make a distinction of order.

The naming of articles: Articles are not always named the same from one edition to the next. For example, the term ‘America’ meant both the country as well as the continent from 1771–1790. In 1842, the article was split into two separate articles, ‘America’ the continent, and ‘United States’ the country. Another example is the article ‘Human’. While in most editions finding the article ‘Human’ is unproblematic. However, in the fifteenth edition of the *Britannica* (1974–2005), there are multiple possibilities to locate the ‘Human’ article. For instance, the 1990 index (v.31, p.831), has the following list: human behaviour, human body, human capital, human culture, human ecology, human fertility, human geography, human intelligence, human migration, human mortality, human nature, human palaeontology, human rights, human sacrifice, humanism. Missing from this list is an individual ‘Human’ article. In cases such as these, the first index entry that has the most references to it is used as the sample text.

Integrated and absent articles: There are many cases in collecting data that the article being looked for did not exist as its own entry. In these cases, if the index includes the word as an entry, then the associated article was used for data analysis. However, only peritextual elements included under the sought-after article were used. For example, an article may be ordered Mammals, Cats, Whiskers. If the sought-after article was ‘Cats’, only the headings ‘Cats’ and ‘Whiskers’ were analyzed and not the parent article.

Index entries without headings: If the sought-after article had an index entry but only directed to a section of an article without a corresponding heading, then no peritextual data was recorded.

Coding was conducted between June 16 and December 8, 2011. A list of the entire corpus is located in the appendix (Tables 6 and 7, pp.170–175).

10.1 Results

10.1.1 Number of headings

	Britannica	Wikipedia
Articles	136	140
¹²⁹ . Headings	14712	3907
Words	47524	9595
AVG words per heading	3.2	2.5

The numbers above display the quantity of headings studied in this analysis. The average word per heading shows that the two encyclopaedias manifest their organization through short headings of two to three words. The main difference between the two is that the *Britannica* is the more verbose one of the two, as shown by the fact that the number of headings of the *Britannica* is approximately 3.75 times larger despite missing four articles. This discrepancy will be explored in the following analysis.

10.1.2 Total heads per article per encyclopaedia

FIGURE 4 (p.179) *Britannica*: Figure 4 displays in descending order, the number of total heads observed in the *Britannica*. As one can see, the article America has over 4.5 times more headings than the next leading article, Canada. This can be explained by the jump in headings between the 1990 and 2005 editions (see Figure 8, p.180). The next three highest articles are Canada, London, and Ocean. While America has 8858 headings, Democracy, Black, Cannon, and Atheism each contain under 200 headings. This difference illustrates that geopolitical entities like America, Canada, and London received substantially more attention from the *Britannica*'s editors.

FIGURE 5 (p.179) *Wikipedia*: The notable trend in the number of heads on Wikipedia is that the gulf between articles is not nearly as vast as the *Britannica*'s. Numerically, the difference between the highest and lowest is only 571 heads on Wikipedia, whereas the *Britannica* has a difference of 8781. In Wikipedia we see that the article Human has replaced America as having the most heads; Democracy moves ahead of Canada; and Ocean, which is fourth highest on the *Britannica* is at the bottom for Wikipedia. While the more geopolitical articles tended to

¹²⁹There are four less articles for the *Britannica* because some editions did not include certain articles.

have the highest number of heads in the *Britannica*, Wikipedia distributed them across the graph. Even though the articles were more evenly distributed, America, London, and Canada continue to stay near the top in both, and Black and Cannon remained at the bottom of both encyclopaedias.

The striking difference between these two graphs is Wikipedia's articles are much closer in number of heads as compared to the *Britannica*'s. The situation is vastly different for the *Britannica* where the article America outpaces every other article by several measures. An explanation for this can be found in the combination of Figures 6 (p.179) and 8 (p.180). If one compares the total 1990 and 2005 heads in Figure 6 to the same years for Canada and America, the two years of those two articles display an unprecedented jump in heads. The reason for this can be explained by the fact that both articles deal with countries and contained an unprecedented amount of detail about their subsequent provinces and states. This choice ballooned the two articles far beyond the other articles studied. However, by 2011, the size of both articles was reduced to levels not seen since 1798. The smaller articles sizes in 2011 can be explained by the fact that instead of retaining the details of both a country and its provinces in the same article, it used hyperlinks to individual articles dedicated to those provinces.

10.1.3 Total heads per edition per encyclopaedia

These two figures display the number of heads used by each encyclopaedia in chronological order. The *Britannica* (Figure 6) exhibits a general trend toward more headings used in each subsequent edition. In the beginning, the editions have relatively few heads, 40 and 180, while the final edition has 707. However, the years of 1842 and 1854 see a substantial drop in heads while 1990 and 2005 see unprecedented increases. What is important to note is that these changes appear to happen in sets of editions. The significant differences between the sets of years can be explained by the fact that the encyclopaedia had different chief editors throughout its lifetime.

For instance, between 1777 and 1797 Colin Macfarquhar was a head editor, corresponding to the 1790-98 editions. For the 1842 edition, Macvey Napier and James Browne were editors and for the 1854 edition it was Thomas Stewart Traill. The relative consistency in the number of heads for each of these editions can be explained by the fact that many articles for the

Eighth edition (1854) were reprinted from the Seventh (Kogan, 1958, p.49). The next span of time which includes the gradual decline for 90 years between 1878 to 1968 can be explained by the fact that the 1878 edition was “the famous ninth” (Kogan, 1958, p.50). Kogan quotes the magazine *Nature* as it extolled that “‘above all other encyclopaedias [the ninth] would be ‘regarded as indicating the highest tidemark of the science, literature and arts of the time’” (1958, p.52). The accolades that the Ninth received made it hard to fix something that was working so well. Therefore, for the next 90 years of the *Britannica* editors relied heavily on reprinting or slightly editing articles from this edition. This situation helps explain the gradual decrease in heads for this 90 year span, as subsequent editions sought to finesse the famous edition. But times changed and so did the needs of *Britannica*’s readers. As has been described in the Section 2 of the analysis, the Fifteenth edition restructured the *Britannica* and was the result of two chief editors J.L. Garvin and Franklin Henry Hooper. Once the Fifteenth edition was published, the *Britannica* switched to a Board of Editors. As seen in the Figure 6, the Board made very few changes from 1974 to 1984 with the exception of adding an index. From 1985 to 2008 the chief editor was Philip W. Goetz, which is reflected by the similarities between the 1990 and 2005 editions. Since the last print edition in 2010 (for which Goetz worked on before he died in 2008), the *Britannica* is once more run by an editorial board. In view of this history, the consistency of the *Britannica*’s heads (or the lack thereof) is a result of the changes enacted by different chief editors.

In Figure 7 (p.179), Wikipedia’s sum of total heads per edition shows that like the *Britannica*, the project started off with only a very few heads. However, this changes by Edition E where instead of falling like the *Britannica*, Wikipedia’s heads increase and then stabilize. Within this trend there is one exception to the gradual change between edition F and G which might be explained by the the increased popularity of Wikipedia in 2005.¹³⁰

Of course the comment that Wikipedia is only 10 years old and the *Britannica* is 240 does limit the direct comparability of the two graphs. However, what must be noted is that the *Britannica* makes substantial structural changes whenever a chief editors decides to do so.

¹³⁰Between Editions F and G (November 13, 2004 to February 16, 2006) there is a dramatic increase in the number of heads. This increase might be explained by the fact that in 2005 “Wikipedia became the most popular reference website on the Internet” (Lefèvre, Jensen & Korsgaard, 2009, p.298). As such, a huge influx of new contributors were active on Wikipedia. However, as it is difficult to determine the actual number of contributors for any article at a given time, this conjecture will have to remain a footnote.

Contrastingly, Wikipedia is a site of mass collaboration and large sweeping changes based on the decisions of an individual rarely result in long standing results. Changes appear to be more gradual from edition to edition, in the same way when the same chief editors are responsible for several editions of the *Britannica*. But of course on Wikipedia there is no chief editor. These changes have been gradually made over the past ten years by thousands of self-organizing contributors.

10.1.4 Number and usage of heads per article per encyclopaedia

As Figure 8 shows, the number of headings differ greatly between articles in the *Britannica*. The shape of the graphs indicates that there is a general unevenness to all articles and there are several instances where continuity is ruptured. What this illustrates is that at different times for different editions, articles were given more or less importance in the encyclopaedia. If we consider that the number and variety of headings is an indicator of the depth of detail and attention given to a particular article, and that such attention is validated by the belief that certain articles are more essential to representing the whole of knowledge, then we have a way of interpreting the graph.

For instance, Cat, Atheism, and Cannon each have editions that do not include these subjects as individual articles. The worst case of omission happens to Black, a subject that disappears completely for one hundred years. Part of this may be due to the fact that the article Black was used as a dictionary root for many early editions. By 1910, it disappears altogether and only resurfaces in rudimentary forms in 2005 and 2011 as a single sentence about “Liturgical Colours”. Additionally, the variety of ranks for these articles are relatively low. As such, these articles can be considered not as essential as America, London, and Canada to describe the whole of knowledge. Alternatively, in some editions the articles Human, Democracy, and Cat are barely more than a cursory description. In other instances, each of these articles have both a variety and a substantial number of heads. Interestingly, Democracy, having never had more than eight heads previously, suddenly shot up past 70 heads in 2005. Explaining why this happened is outside of the scope of the thesis, but what can be accurately described is that the editors saw that Democracy has now become a much more essential subject to understanding the world.

On the whole, the headings per article of the *Britannica* show that there has been a general

uneven attention given to each article over the span of the encyclopaedia's lifetime. Again, like the comparison of Figures 6 and 7, the unevenness of the *Britannica* contrasts significantly with the much more gradual changes in the number and variety of Wikipedia's heads. Like the *Britannica*, each article starts out with relatively few headings. However, Wikipedia's rank usage gradually increases until it more or less plateaus in each article. There are a few exceptions like Canada, America, and London that see a short descent before plateauing. These results are quite different than *Britannica*'s. The consistency of the plateaus shows that across Wikipedia, there is an effort to equalize the number of ranks used between and within articles. As well, none of these articles have been omitted from Wikipedia for any period of time. As such, each article has been deemed to be important to the goal of representing the whole of knowledge. However, the number of headings used for each article indicates that not every article is equally essential. Black, Ocean, and Cannon all plateau around the 20 mark; Canada, Atheism, and Democracy between 40 and 50; London, America, and Human around 60 and 70. Given these results, there appears to be a trend for Wikipedians to confidently describe an article within 20, 40, or 60 head increments. Further research is required to identify whether such a trend is global across Wikipedia.

10.1.5 Percentage of ranks used per edition per encyclopaedia

FIGURE 10 (p.182) *Britannica*: For the first four editions of the *Britannica*, Rank 2 dominates. This begins to lessen in 1842 and 1853, giving additional weight to Ranks 3 and 5. We see further transitions of reliance given to Rank 4 in 1878 and then to Rank 5 in 1910. After this point, Ranks 6 and 7 play a much more crucial aspect of organizing the contents of the *Britannica*. However, the 2011 online version returns dominance to Ranks 3, 4, and 5, but with a similar distribution of the heads.

FIGURE 11 (p.182) *Wikipedia*: The first three editions of Wikipedia share a similar distribution percentage between Ranks 1, 2, and 3. Like a waning moon, reliance on Rank 1 lessens until the G edition, where it stays relatively stable as a mere sliver of percentage. Similarly, Rank 2 drops to 19-20% in edition H. Rank 3 stabilizes between 35-40% by G edition. By H edition, the distribution of ranks is relatively stable.

It is when we compare the distribution of heads of each encyclopaedia that degrees of

similarity truly begin to emerge from the data. What we see is that with each subsequent edition, the distribution of rank usage becomes increasingly complex. For the first few editions, both encyclopaedias keep to only a few number of ranks. Then, by the fifth edition, both encyclopaedias start a process of diversification, where the majority of headings are relegated to three ranks that make up about one third to a quarter each. The remaining top and bottom ranks are then distributed to the final quarter.

What can be assessed from these results, is that despite the discrepancies between the number of heads used, each encyclopaedia creates a diamond shaped hierarchy. The Ranks 1 and 2 being significantly less numerous than ranks 3, 4, and 5. After this bulge, instances of ranks 6 and 7 are relatively small. Therefore the history of the distribution of ranks represents a systematic similarity between the two encyclopaedias.

10.2 Common and unique heading words

As the the previous tables and diagrams have shown, the number of headings and the type of headings have changed throughout time and between the two different encyclopaedias. While these quantitative differences are instructive in showing how the encyclopaedias have implemented their structure at the level of the article, they offer little commentary on what words or topics are used to systematically organize each encyclopaedia. Examining what words are common to organizing these two encyclopaedias will “reveal even more about the way discourse is intimately linked not only to knowledge and theory, but to epistemology — beliefs about what can be known, how it can be known, in what form it can be expressed and how it should be argued” (Bazerman, 1988, p.174). The methodology of this analysis includes a control corpus (*Britannica*) that will show how encyclopaedism traditionally textualizes its epistemology. Similarities between the two encyclopaedias will display how encyclopaedism has retained its communicative power to define meaning on Wikipedia. Comparatively, differences between the *Britannica* and Wikipedia will show how the epistemic values of wikis have exercised communication power in interfering with Wikipedia’s encyclopaedism at the textual level. Therefore the current study establishes two sets of categories; words that are common to both encyclopedias and words that are unique in each.

10.3 Most influential heading words

10.3.1 Most frequent words common to both encyclopaedias

TABLES 10 & 11 (p.183): These two tables show the most frequent words used (excluding conjunctions and prepositions) in all the headings examined of both encyclopaedias. Table 10 sorts the words by most frequent in the *Britannica* and Table 11 sorts the words by most frequent on Wikipedia.

10.3.2 The most frequently used words that are common to both encyclopaedias

FIGURE 12 (p.184):¹³¹ This figure highlights the top most frequent words used by both the *Britannica* and Wikipedia. While these words are common to both, each encyclopaedia uses these words with different frequencies. This diagram displays these differences by ranking the most frequent word to the least frequent word of each encyclopaedia, and calculating this difference to find out how similar the use of these words are. For instance, a longer purple bar indicates that the word, although common to both, is significantly used in the *Britannica*. Likewise, a long blue bar indicates that Wikipedia uses the word more often than the *Britannica*. Drawing attention to the middle of the graph, one sees that there is almost no difference between the use in either encyclopaedia, indicating that these words are of equal high frequency in both encyclopaedias. Highlighted in yellow are the central thirteen words that will be investigated for structural similarity.

Interestingly, the full list shares similarities with Diderot's plan for the *Encyclopédie*. For instance, history figures prominently in his plan under 'Memory'. Under 'Reason' we find *religion, politics, economics, geography, medicine* and *hygiene* (which can be correlated to *health*), and *military architecture* and *tactics* (elements of war). The categories of *climate* and *world* can be found within a combination of Diderot's categories, but they are not explicitly described. Missing from Diderot's plan are the categories of *education, society, and culture* that are prominent within *Britannica* and Wikipedia. Thus, even before the rest of the analysis has

¹³¹ A caveat to this is that the graph contains two sets of synonyms, *culture/cultural* and *society/social*. In the *Britannica*, the words *culture* and *society* were used more often than *cultural* and *social*, while for Wikipedia, the opposite is true. The graphs shows that when these synonyms are combined, there is actually very little difference in usage by either encyclopaedia. It is for this reason that these two sets of synonyms are combined for the purposes of this study. Another note to add is that the words *world* and *war* are often (but not always) seen together as relating to the two World Wars of the twentieth century.

happened, there are signs that Bacon's outline of knowledge continues to greatly influence the structure of encyclopaedism.

10.3.3 Most frequent words unique to each encyclopaedia

TABLES 12 & 13 (p.185): These two tables show the most frequent words in each encyclopaedia that do not exist in the headings of the other encyclopaedia. Looking at the differences between the unique words of each encyclopaedia is quite intriguing. In the *Britannica*, we see words that are objects, like *land*, *animal*, *plant*, and *relief* (the geographic term used for elevation). There are also governmental terms like *administration* and *welfare*, and economic terms like *industry*, *resources*, *growth* (as in economic growth) and *conditions* (as in economic conditions).

The unique Wikipedia words differ in that some are more subjects than objects: *etymology*, *biology*, *philosophy*, *physiology*, *sexuality*, *psychology*, and *spirituality*. *Varieties* is a term that is usually associated in conjunction with breeds or species. What stands out between the two of these lists is that while many of the unique terms of the *Britannica* describe the external physical world, Wikipedia's unique words describe the worlds of the self and society. This fact will be espoused upon in the discussion section of the thesis.

10.3.4 Multiple frequencies

While a straight forward frequency is an excellent indicator of what is important, it is by no means the only way of observing importance. This is especially true when it comes to analyzing headings. To understand the influence that these words have on the text, five different frequencies have been observed: rank, sequence, command, consistency, and longevity. These additions describe a more complete picture of how certain words organize an encyclopaedia. Each of the most frequent words are analyzed for these qualities, their totals are tallied, and the final scores represent not only the most frequent, but also the words that do the heavy lifting in systematically organizing the encyclopaedias.

For analysis, the first quality to look for is the repetition of certain words within headings. If a word is repeated throughout different articles, editions, and encyclopaedias, this may be a sign that it forms an essential part of understanding encyclopaedic topics. The repetition of such categories forms a familiarity between the reader and a text. As Kools *et al.* describe,

readers have expectations about how a particular kind of text should be organized (2008, p.834). These expectations arise from previous experiences with texts of the same kind. As such, an encyclopaedia that is interested in providing quick access to the text will repeat headings to create a familiarity which “in turn makes the text easier to follow and its organization easier to grasp” (Kools *et al.*, 2008, p.834). It is in this way that the frequency of such headings will offer insight into what categories are important to the organization of encyclopaedias.

Rank: TABLE 14 (p.186): As has been discussed throughout the methodology, rank and hierarchy play an important role in how a reader receives a text. In the corpus, each word has a relative rank from the beginning of the article. These ranks have been counted for each instance of the common and unique words. In order to measure such importance, a fractional scale has been used to weight a given rank. For example, the frequency of a heading at level 1 is multiplied by 1, rank 2:1/2, rank 3:1/3, and so on. These numbers are then added together as a word’s Rank Score.

Sequence: FIGURE 13 (p.187): Sequence plays an important role in structuring an article. Traditionally, those aspects that are placed at the beginning of an article correspond to the most essential qualities of a text. Williams & Spyridakis reiterate this when they state that “the more important, more inclusive, ideas are presented higher in the content structure and the less important, supporting details are presented lower in the content structure” (1992, p.64). Therefore, if a heading word is often used at the beginning of an article, it can be considered more essential than a heading word used at the end. In order to measure for sequence, the position of a heading word is attributed to a percentile of the total length of the article. Then, instances of this word have been grouped into increments of ten per cent (10-19%, 20-29%, 30-39%, etc.). As with rank, a degree of weight is added to each position. Thus, positions in the first ten per cent are weighted heavier than those in the final ten percent. The sequence score is the multiplication of frequency of a word at a given position multiplied by a decreasing weighting metric: 1(0-9%), 0.9(10-19%), 0.8(20-29%), etc.

Command: FIGURE 15 (p.189): Command is the number of “children” that a “parent” heading has directly beneath it. The assumption is that the more subheads that a head contains, the more important it is. In the case of creating a score, the frequency of a heading is divided by the number of subheads. This score is also augmented by how often a head has at least one subhead. If it does not have any subheads, it is less effective in commanding the article. The final score is achieved by comparing these two aspects.

Consistency: FIGURE 16 (p.189): Consistency is used to see if a certain heading word is used across multiple kinds of articles. As the word appears in more kinds of articles, that word becomes an essential word for understanding all subjects. The score for this quality is achieved by counting the presence of a word in each of the ten articles.

In terms of consistency, not one word in either encyclopaedia in the common or unique lists scored perfectly. The word that is most common across all of the articles is Wikipedia’s *etymology*, which does not occur in either Democracy or Ocean. The rest of Wikipedia’s unique words are not consistently used. *Britannica*’s unique words fair better, but barely. Only a few words are used in 50% of the articles studied. The common words do a better job of being consistent across articles. Overall, out of a possible 130 (10 articles per word), the *Britannica* scored 59 and Wikipedia 63. In terms of percentage, this means that these common words only appeared in 45% and 48% in the *Britannica*’s and Wikipedia’s respective articles.

Longevity: FIGURE 17 (p.189): Finally, the last quality is how long has the word been in use. If a word has been used for a longer period of time, it is considered more essential to understanding a topic. Like consistency, this score is achieved by counting the presence of a word in each of the fourteen editions. Figure 17 shows the number of editions that have at least one instance of that word in a heading. Of Wikipedia and the *Britannica*, only *history* has been in all 14 editions of Wikipedia. After that, *government*, *climate*, and *history* appear in 13 editions of the *Britannica*. Calculating the total scores of each encyclopaedia, these common words appear in 73% of the *Britannica*’s editions and 76% of Wikipedia’s. The unique words appear in slightly less editions, 69% for the *Britannica* and 64% for Wikipedia. As such, the words for both encyclopaedias are comparable in terms of longevity.

10.3.5 Comparison of most influential common and unique words

FIGURES 19 & 20 (p.193): After tallying up the scores of each of the five qualities (Figure 18, p.192) the resulting figures show the words that are essential to the organization of the articles analyzed. The transposition of the common words indicates that certain words are more or less important to describing the essential qualities of encyclopaedic texts. The reasons for such movements are an obvious direction for future research. While positions may have changed, it must be recognized that the common words listed here show that there is a substantial similarity of concepts that organize the two encyclopaedias. Here we see that among the common words, there have been changes in terms of the importance that certain words play. For instance, in *Britannica* the most influential word is *cultural/culture*. In Wikipedia, this position is held by *history*. If we observe the top six words, there are a number of words that are common. For instance, *cultural/culture*, *history*, *government*, *war*, and *climate* are all seen in the top six, all but *climate* transposing within these positions. Of those that are both in the top six and the bottom seven, we see the curious switch of *social/society* from position 4 in the *Britannica* to position 9 in Wikipedia and *geography* in position 9 in the *Britannica* to position 4 in Wikipedia. Beyond those two, *education*, *religion*, *health*, and *world* switch positions. In addition to *climate*, both *economy* and *political* retain their positions in both encyclopaedias.

10.4 Conclusions

Number of headings: To generalize, the results of studying ten articles from two encyclopaedias show that there are substantially strong correlations in their organizational structures. Both encyclopaedias show that for most articles there is a period where there are relatively few heads used. While the number of total heads may vary, the number of ranks tends to increase in most articles, showing that for these articles, time grants the article a greater level of complexity.

Additionally, if one examines the progression of heading use from edition to edition, Wikipedia and *Britannica* reveal how they have a structural isomorphology. As Figures 10 and 11 (Percentage of ranks used within each edition) show, the two encyclopaedias started out with using only a very small set of ranks. The history progressed into a much more complex rank structure which can be described as a diamond shape. The first and last ranks were used

sparingly while the middle ranks were used extensively. This similarity is striking for the fact that even though the development of articles has been achieved through contrasting editorial authority, the overall approach to how articles should be structured is closely related.

While these results show that the two encyclopaedias are similar in the complexity of their structure, the history of editions shows a different story. Wikipedia displays a resoluteness in the articles studied. This is not the case for the *Britannica* which completely omitted or reduced articles to cursory descriptions. Another difference is that the *Britannica* tends to have much longer articles and inconsistent heading use between editions. Wikipedia on the other hand tends to be shorter in length and has greater rank consistency.

What these trends show is that the authority of the editors has a marked difference on the outcome of the articles themselves. In the case of the *Britannica*, editorial decisions were often made by one or two chief editors. As a result of this centralized authority, the editors were able to impress upon the encyclopaedia their individual ideas as to what the encyclopaedia should cover and how much should be dedicated. The idiosyncrasies of the editors can thus be seen by which articles are included, omitted, deeply detailed, or are cursory summaries. Figure 6 displays how authoritative editorial decisions lead to a general unevenness from edition to edition. The contrast in this regard to Wikipedia's usage of heads is strong and obvious. Because there is not a central authority to the encyclopaedia, the whims of individuals can only be actualized if they are mutually accepted by the community. As the various editions chosen for the study span the course of about a year, the typical back and forth of edit wars and incremental changes that Viégas, Wattenberg & Kushal (2004) report take a backseat to the overall movement of the article. Over time those conflicts are resolved and the resulting contributions that remain represent the gradual accumulation of consensus, which translates directly into the growth of an article's heading structure. In this way we can see that the differences in editorial authority has resulted in dramatically different histories of article progression. Returning to the overview of contexts described in Section 2 of the analysis, this finding is in keeping with the view that the status of encyclopaedic authorship is a non-discriminative strategy. While this difference results in a different method of production, it does not affect the generic membership to the genre of encyclopaedism.

Kinds of categories: Despite the changes of position in Figure 19 (p.193), the very notion that these words have such comparable results indicates that there is a tradition of organization that is being followed by Wikipedia. As has been previously indicated, these categories are in keeping with a number of categories outlined by Bacon's outline of knowledge.

In terms of difference, the number of headings and number of ranks is quite large between the encyclopaedias. However, if one looks at the distribution of heads and the distribution of ranks within articles that difference begins to disappear. Looking to sequence, we see that certain words tend to appear in similar places at comparable intensities. For instance, Figure 14 (p.188) showed similarities in sequence in *climate*, *economy*, *geography*, *world*, *war*, *social/society*, and *government*. In terms of rank distribution, Figures 10 & 11 (p.182) showed that both encyclopaedias followed a pattern of limited heading use to more complex heading use. These findings point to the fact there is an implicit understanding by the editors of both encyclopaedias that there is a conventional place for these topics to begin structuring an article.

One observation that must be made is that a great number of the common words are geopolitical in nature. This is partly due to the fact that 30 per cent of the articles studied are geopolitical entities (Canada, America, London). As such, the results may be seen as skewed in favour of heads that are typical of geopolitical articles and not extendable to other unrelated topics. In reviewing the consistency of the common words, eight of the thirteen words were only in 50 per cent of the articles. This indicates that while geopolitical words may have overwhelmed the results, there were a number of other articles that included those words. In view of this, future research would benefit by examining whether this list of thirteen words are also frequently used in non-geopolitical articles.

Despite the similarities previously described, there are two outstanding features within the common words that describe a distinguishable difference between the *Britannica* and Wikipedia. Looking at Figure 14:History (p.188), we see Wikipedia greatly exceeds the number of heads used by the *Britannica*.¹³² This is supported by the fact that not only is *history* frequently used, it is used mostly at the beginning of Wikipedia articles, giving it an unchallenged right to define Wikipedia's articles. Additionally, *etymology* (the history of words), is similarly used in

¹³²Only in two other words *political* and *religion* do we see Wikipedia override the *Britannica* number of heads (but their peaks are only around ten heads high).

abundance at the start of articles. What this indicates is that Wikipedia is emphatically a space for history.

On the other hand, there is the discrepancy described by the unique words that each encyclopaedia uses. In this regard, there has been a real shift in the kinds of topics used to describe the articles studied. The discussion section of the thesis will go into detail as to why this might be occurring.

In view of the results, the textual structure of articles is directed by the epistemic values of systematic organization and consistency to represent the whole of knowledge. As has been mentioned, the common words used to organize both Wikipedia and the *Britannica* shared a number of similarities to Diderot & d'Alembert's (and therefore Bacon's) outline of knowledge. As such, the study concludes that given the frequencies recorded of the common words, not only is Wikipedia's article structure similar to the *Britannica*'s, but it also shares commonalities with categories of knowledge that are rooted in seventeenth century philosophies of knowledge.

Part V

Discussion

At the end of the twentieth century, society was beginning to dramatically transform. As described by Castells, this transformation was established through the development of robust and decentralized communication infrastructures, the globalization of flexible flows of capital, and the necessity for individual freedoms and open communication (2001, p.2). It is through the confluence of these processes that the network society forms the boundaries of the social relationships that we have with one another and the world. As such, it is through the logic of the network that we communicate with one another and construct cultural meanings. Castells argues that when two networks come into conflict, the adjacent networks attempt to reprogram each other's values and goals by controlling discourse through communication power. What is at stake in network conflicts is therefore the ability to contribute to the making and interpretation of one's own culture. Therefore knowledge (as well as all other cultural meanings) in the network society are transformed by the meshing of conflicting network values. When we visit Wikipedia to refer to something unfamiliar, we are reading information that has been shaped

by this meshing of conflicting discursive activities. The current thesis has outlined that these activities are guided by two sets of epistemic values. In effect, the main question of the thesis “What is the epistemology of Wikipedia?” has been addressed: Wikipedia’s epistemology can be considered to be determined by wiki and encyclopaedic epistemic values.

Additionally, the thesis concluded that the encyclopaedic values of systematic organization and consistency were found to be prevalent in the headings that structure English Wikipedia’s articles. In fact, it was discovered that the most frequent categories Wikipedia used were found to be significantly similar to the *Britannica*’s. Upon further reflection, both sets of frequent categories were considered to have ties to Diderot’s “System of Human Knowledge” which was inspired by Francis Bacon’s *Instauratio Magna* (1620). In this respect, Wikipedia’s production of knowledge can be seen to be entangled in both seventeenth century ideas about knowledge and the two and a half centuries of the *Britannica*’s encyclopaedic development. Taking this into consideration, the world that Wikipedia presents in its articles is one that is restricted by traditionally Western-based categories of knowledge. With this in mind, even though the publishing platform may be inclusive in terms of infrastructure, there is an epistemological exclusivity at work, one that challenges the very openness that has made Wikipedia such a valuable source of knowledge.

These findings have been arrived at by addressing the joint research question of “What is the genre and epistemology of encyclopaedism?”. The genre of encyclopaedism was defined through its communicative purposes and its epistemology was revealed through the genre’s recurrent social contexts. In this light, Wikipedia was shown to exhibit the same strategies of achieving the communicative purposes of encyclopaedism. Furthermore, two epistemic values were found to structure Wikipedia’s articles. While these findings are important, the picture remains incomplete. What is left unanswered is how the remaining epistemic values manifest themselves within the text of Wikipedia and how they effectively shape its knowledge. Likewise, there are more theoretical and methodological limitations to the thesis that prevent the broader implications from being thoroughly analyzed.

10.5 Theoretical and Methodological Limitations

One of the aspects of the current research that could not be completed was a direct comparison of the epistemological values of wikis and encyclopaedism. This resulted from the fact that each set of values was not equally presented in the literature review. Arriving at the end, the current thesis only describes the two sets of epistemic values, as opposed to illustrating how these two sets interact with one another. It is hoped that given this new description, future research will examine these epistemological interactions. Supporting this program of research has been the philosophies of technology from Latour and Castells. These two scholars view networks as heterogenous assemblages of diverse actants joining, combining, and distributing different values, goals and purposes. As illustrated by the current thesis, the sort of questions asked by this theory seek to understand the structure of society and how meaning is mediated through its technological networks. As such, the current thesis took on a descriptive quality, describing Wikipedia in terms of its parts and functions. Contrastingly, this approach has been limited in its ability to be critical of the situation described.

Bearing this in mind, future directions for the current research could involve critically analyzing the hegemonic power structures that are dominating the outcomes of conflicts over values and goals. While the inclusion of Heidegger in the current thesis is a nod to critical theory, there are other perspectives (like those of Marx (1846), Hall (1980), and Habermas (1954)) that could provide a critique of Wikipedia's epistemological foundations. Positioning Wikipedia within these perspectives could problematize the hegemonic power structures that are entwined with the epistemologies of capitalism, institutionalized racism and sexism, and political control. In other words, Wikipedia's epistemology could be examined in terms of post-colonial theories of power by questioning how the making of meaning is constrained by particular colonial mobilizations of knowledge, even in light of its flexible, open, and inclusive characteristics.

Additionally, this thesis includes some methodological limitations that must be considered. While Bazerman's four recurrent social contexts are used to examine encyclopaedism, they were not the only set available. Within the literature there are number of other categories that have been used to understand genres. For instance, Levine (2007) offers five constraints (production

constraints, the production environment, production routines and practices, the production of characters and stories, the role of the audience in production). Martin (1992) understands meaning in discourse by analyzing the components of its register, as explained by Bawarshi (2000) “what actually takes place communicatively (the ‘field’), who is taking part (the ‘tenor’), and what role language is playing (the ‘mode’)” (p.350). Looking at just these two alternative contexts, despite their different trajectories, they have some aspects in common with Bazerman’s lexicon, connected literature, audiences, and authors. Bazerman’s methodology was specifically chosen over alternatives due to the similarity in time frame (eighteenth century to present) and the nature of the research object (both the scientific article and encyclopaedism are dedicated to questions about knowledge). In many regards, this methodology proved appropriate to satisfy the questions of the thesis. However, through the course of analysis Bazerman’s isolated concepts of audience and authors posed a theoretical incongruity with Castells’ argument that the network society blurs those two roles. The solution chosen was to continue to apply Bazerman’s contexts and address these issues within the discussion of the contexts themselves. Similar research efforts may benefit from a different conceptualization of the social contexts.

In terms of the textual analysis, the textual space of article pages were examined to investigate how an epistemic value surfaces within Wikipedia. This is but one of many possible spaces that could have been analyzed. Examining to what degree encyclopaedism is present in the history pages, talk pages, or user pages of Wikipedia would offer different examples of how encyclopaedism constructs meaning within Wikipedia. Additionally, while the thesis analyzed heading structure, other units could have been used. For instance, citation sources could have been compared or a rhetorical analysis of the move structure within articles could have been equally revealing of epistemological choices in action.

10.6 Summary of Contributions

The value of the current thesis is that it is the first step in understanding how knowledge in the network society is transformed by conflicting networks. This process entailed revealing the epistemic values of Wikipedia’s conflicting networks, wikis and encyclopaedism. For wiki values, the study relied heavily on Ruth & Houghton’s (2009) epistemology of wikis. To add weight to this list, a number of studies were used to support their claims. As such, the wiki

epistemology is characterized by the values of collaboration, co-construction, cooperation, trust in the community, and constructionism. The opposing network, encyclopaedism, posed certain historical problems in terms of establishing a set of epistemic values. As such, the thesis' major contribution to understanding the epistemology of Wikipedia comes in the form of identifying encyclopaedism as a genre with five unique epistemic values: utility, systematic organization, consistency, authority, and trust in experts.

Additionally, in defining encyclopaedism in terms of a genre with discriminative strategies to achieve specific communicative purposes, the research has provided a framework in which to remedy some of the confusion surrounding Wikipedia. For instance, the discriminative strategies can be used to resolve Emigh & Herring's question: "How is it that the wide-open participation structure of a wiki can reproduce traditional print norms?" (2005, p.9). While their study and the current study came to similar results in terms of the similarity between Wikipedia and traditional encyclopaedias, Emigh & Herring framed it as an incompatibility. However, given the chapter on genre and the descriptions of discriminative strategies, the similarities between *Britannica* and Wikipedia's article structure should not be surprising. If Wikipedia considers itself to be an encyclopaedia, it must abide by the discriminative strategies of the genre, which includes consistency and systematic organization. As such, Wikipedia can be seen to behave similarly to older encyclopaedias because it identifies itself as an encyclopaedia. To not follow these discriminative strategies would deny it generic membership. In other words, Wikipedia would cease to be an encyclopaedia if it stopped using systematically organized and consistent content.

Emigh & Herring's question also addresses a different problem. It assumes that "wide-open participation" runs counter to the discriminative strategies of encyclopaedism. The confusion on this point is that the discriminative strategy of trust in experts has been confused with the tactic of using expert authors. Section 2 of the analysis argues that encyclopaedic authors have not been constrained to this role. Instead they have carried a spectrum of authorship, including editors, generalists, and volunteers that are not necessarily experts on the topics they write about. Therefore the issue that Emigh & Herring bring to the fore is that the expertise of authors is wrongly assumed to be a discriminative strategy of encyclopaedias. It is instead, part of the evolutionary milieu, a malleable characteristic that depends on what an individual encyclopaedia is trying to achieve beyond being an encyclopaedia. However, what *is* necessary

of all encyclopaedias is that they must verify their knowledge by pointing to an expert source. The tactics for this strategy vary and include the use of citation, as is heavily utilized by Wikipedia, or expert authors as in the case of the *Britannica*. As this explanation of Emigh & Herring's question displays, the use of discriminative strategies revealed in this study are pivotal to understanding what Wikipedia is and how it functions.

10.7 Future Research

While revealing the epistemic values of encyclopaedism (also known as the discriminative strategies of the genre) has been the primary goal of this thesis, it is only to enable a larger research project to take place. This broader project asks how the epistemic conflicts have resulted in a third and dynamic epistemology of Wikipedia. Ultimately this research would provide insights into how knowledge is transformed in the network society. Future research can thus use the two sets of epistemic values described in this thesis and compare how each one has been communicated in different spaces within Wikipedia, illuminating the topography of Wikipedia's epistemology.

In view of this, the literature and analysis of the current thesis has indicated some potential sites of investigation. In the analysis of textual structures, Wikipedia used the word *history* as a heading far more than the *Britannica*. Why has this happened? Could this be the result of an epistemic conflict? Additionally, the textual analysis observed a list of heading words not used by the *Britannica*. This list of words (*sexuality, psychology, philosophy, civilization, gender, and spirituality*) are all categories used to describe how people fit within and interact with society. In other words, these topics could be seen to describe articles in terms of self-identification, a wiki epistemic value. Whether this is true or not will have to be left up to future research to discover.

Other avenues of research concern how the epistemic values of wikis and encyclopaedism conflict with one another in different situations. For instance, are the values of authority and self-identification the grounds that the Userbox controversy was based upon (Westerman, 2009)? Can the conflict between the two groups of Wikipedians, the inclusionists and the deletionists,¹³³ be explained by a conflict between wiki and encyclopaedic values? Further to the point, can the rhetoric about Wikipedians being "amateurs" be explained by an attempt by encyclopaedic

¹³³The *Economist* (2008) and Kostakis (2010).

values to reprogram a wiki space?

With the epistemic values of encyclopaedism explained in the current thesis, questions like these now have a foundation to build from. As such, studies that continue this research may not only describe the shifting values of Wikipedia's epistemology but also explain how knowledge is transformed and produced in the network society.

10.8 Conclusion

Reaching the end of this thesis, what has been realized through the extensive research about encyclopaedism is that it has been a model project of innovation, collaboration, and tenacity. The sheer magnitude of its goal and the unrelenting centuries of its pursuit speak volumes about the passion and imagination of its millions of contributors. However, within the context of the network society, the conflicts between encyclopaedism and wiki ultimately pose some very real problems for Wikipedia's sustainability as a source of knowledge in the twenty-first century.

As shown by the textual analysis, encyclopaedism is a very powerful network in defining discourse. As such, if future research on this matter shows that there is an imbalance of power between encyclopaedism and wiki, then the very values that make Wikipedia a unique, original, and model project of the network society could be subsumed by the communication power of encyclopaedism. If this is the case, then the epistemic values of Wikipedia will seek only to recapture a golden age of expert authority and not represent the collaborative and flexible values of the network society.

Perhaps it is beneficial to rephrase this situation in the context that encyclopaedism emerged from. The eighteenth century encyclopaedias did not succeed solely on the technical advances of the printing press nor the creation of new social roles in encyclopaedic corporations. They succeeded because these innovations were combined with a knowledge system that made sense for the time; it reflected the intellectual transformations occurring within society.

As we have seen in our own time, the network society has similarly emerged from new technological affordances that allow information and social roles to flow effortlessly. That being said, retaining a perspective that knowledge continues to have a static and singular truth lacks the spirit of the network society. The consequences of this statement is that Wikipedia will eventually lose its societal relevancy if it allows encyclopaedic values to overwhelm episte-

mological conditions. If this prediction is to be avoided, then Wikipedia needs to reinvigorate its commitment to wiki values, to judge knowledge, as Lévy explicates, not for its truth but for its reflexivity (2011, p.332). Above all else, this thesis has witnessed how the success of encyclopaedias was achieved by embracing new technologies, new social roles, *and* new epistemologies. If this is what it means to collect and share knowledge in the network society, then this tradition is not only aptly suited to the needs of today, but must be followed to the letter.

Part VI

End Matter

11 Bibliography

References

- Alexa.com (Retrieved April 25, 2013). Wikipedia.org: Traffic stats: Reach%. <http://www.alexacom/siteinfo/wikipedia.org>.
- Alexa.com (Retrieved December 27, 2012a). Alexa top 500 global sites. <http://www.alexacom/topsites>.
- Alexa.com (Retrieved July 1, 2012b). britannica.com. http://www.alexacom/search?q=Britannica.com&r=topsites_index&p=bigtop.
- Allen, G. (2000). *Intertextuality*. Routledge, London, UK; New York, NY.
- Altman, R. (1987). *The American film musical*. Indiana University Press, Bloomington, IN.
- American Psychological Association (2010). *Publication manual of the American Psychological Association*. American Psychological Association, Washington, DC.
- Andersen, J. (2002). Materiality of works: the bibliographic record as text. *Cataloging & classification quarterly*, 33(3-4):39–65.
- Antin, J. and Cheshire, C. (2010). Readers are not free-riders: Reading as a form of participation on Wikipedia. In *Computer Supported Cooperative Work (CSCW) 2010*, pages 1–4. ACM, Savannah, GA.
- Antin, J., Yee, R., Cheshire, C., and Nov, O. (2011). Gender differences in wikipedia editing. In *WikiSym'11*. ACM.
- Ashmore, H. S. (1962). Editing the universal encyclopedia. *American Behavioral Scientist*, 6:15–18.
- Askehave, I. and Nielsen, A. (2005). Digital genres: a challenge to traditional genre theory. *Information Technology & People*, 18(2):120–141.
- Askehave, I. and Swales, J. M. (2001). Genre identification and communicative purpose: A problem and a possible solution. *Applied Linguistics*, 22(2):195–212.
- Baran, P. (1964). On distributed communications: Introduction to distributed communications networks. Memo, The Rand Corporation.
- Barney, D. (2004). *The network society*. Polity, Cambridge, UK; Malden, MA.
- Bawarshi, A. (2000). The genre function. *College English*, 62(3):pp. 335–360.
- Bazerman, C. (1988). *Shaping written knowledge: the genre and activity of the experimental article in science*. University of Wisconsin Press, Madison, WI.

- Benkler, Y. (2006). *The Wealth of Networks*. Yale University Press, New Haven, CT; London, UK.
- Bentham, J. (1776). *A Fragment on Government: Being an Examination of what is Delivered, on the Subject of Government in General, in the Introduction to Sir William Blackstone's Commentaries: with a Preface, in which is Given a Critique on the Work at Large*. T. Payne; P. Elmsly; and E. Brooke.
- Bentham, J. (1823). *An introduction to the principles of morals and legislation*, volume 1 of *An Introduction to the Principles of Morals and Legislation*. Printed for W. Pickering.
- Berkenkotter, C. and Huckin, T. N. (1995). *Genre Knowledge in Disciplinary Communication: Cognition, Culture, Power*. Lawrence Erlbaum Associates, Hillsdale, NJ.
- Berners-Lee, T. (1989). Information management: A proposal. <http://www.nic.funet.fi/index/FUNET/history/internet/w3c/proposal.html>.
- Berners-Lee, T. and Connolly, D. (1993). Hypertext markup language (HTML): A representation of textual information and metainformation for retrieval and interchange. Technical report, IIR Working Group.
- Beschastnikh, I., Kriplean, T., and McDonald, D. (2008). Wikipedian self-governance in action: Motivating the policy lens. In *Proceedings of International Conference on Weblogs and Social Media*, pages 27–35, New York, NY. Association for the Advancement of Artificial Intelligence.
- Bhatia, V. (1993). *Analysing genre: language use in professional settings*. Applied linguistics and language study. Longman, London, UK; New York, NY.
- Bhatia, V. (2004). *Worlds of written discourse: A genre-based view*. Continuum, London, UK; New York, NY.
- Biuk-Aghai, R. P. and Lei, K. H. (2010). Chatting in the wiki: synchronous-asynchronous integration. In *WikiSym '10: Proceedings of the 6th International Symposium on Wikis and Open Collaboration*, pages 1–2, New York, NY. ACM.
- Bos, A. (1989). *Cosmic and meta-cosmic theology in Aristotle's lost dialogues*. Brill, Leiden, NY.
- Brazzeal, B. (2011). Citations to Wikipedia in chemistry journals: A preliminary study. *Issues in Science and Technology Librarianship*, pages <http://www.istl.org/11-fall/refereed2.html>.
- Bringhurst, R. (2004). *The Elements of Typographic Style*. Hartley & Marks Publishers, Point Roberts, WA.
- Bryant, S. L., Forte, A., and Bruckman, A. (2005). Becoming Wikipedian: Transformation of participation in a collaborative online encyclopedia. In *GROUP '05: Proceedings of the 2005 International ACM SIGGROUP conference on Supporting group work*, pages 1–10, New York, NY. ACM.

- Buringh, Eltjo; van Zanden, J. L. (2009). Charting the “Rise of the West”: Manuscripts and printed books in Europe, a long-term perspective from the sixteenth through eighteenth centuries. *The Journal of Economic History*, 69(2):409–445.
- Burke, P. (2000). *A social history of knowledge: from Gutenberg to Diderot*. Polity Press & Blackwell Publishers, Cambridge, UK; Malden, MA.
- Burke, P. (2002). Coping with Gutenberg: The information explosion in early modern Europe. *Instituto de Estudos Avançados da Universidade de São Paulo*, 44.
- Burke, P. (2010). Peter Burke on the social history of knowledge, 1750–2000. <http://theoryculturesociety.blogspot.com/2010/11/interview-with-peter-burke-on-social.html>.
- Burke, P. (2012). *A Social History of Knowledge II: From the Encyclopaedia to Wikipedia*. Polity, Cambridge, UK; Malden, MA.
- Butler, B., Joyce, E., and Pike, J. (2008). Don’t look now, but we’ve created a bureaucracy: the nature and roles of policies and rules in Wikipedia. In *CHI ’08: Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems*, pages 1101–1110, New York, NY. ACM.
- Byars, J. and Meehan, E. R. (2000). *Television: the critical view*, chapter Once in a lifetime: Constructing “The Working Woman” through Cable Narrowcasting, pages 144–168. Oxford University Press, New York.
- Carnall, M. (2009). Chasing the online audience. In *UMACJ 2/2009*, pages 37–40.
- Castells, M. (1996). *The Rise of The Network Society*. Wiley-Blackwell, Chichester, UK; Malden, MA.
- Castells, M. (1997 [2010]). *The Power of Identity: The Information Age: Economy, Society, and Culture*. Number v. 2 in Information Age Series. Wiley.
- Castells, M. (2000). Toward a sociology of the network society. *Contemporary sociology*, 29(5):693–699.
- Castells, M. (2001). *The Internet Galaxy: Reflections On The Internet, Business, and Society*. Oxford University Press, Oxford, UK; New York, NY.
- Castells, M. (2004). *The Network Society: A Cross-Cultural Perspective*, chapter Informationalism, networks, and the network society: a theoretical blueprint, pages 3–48. Edward Elgar Publishing, Cheltenham, UK.
- Castells, M. (2009). *Communication Power*. Oxford University Press, Oxford, UK.
- Castells, M., Cardoso, G., et al. (2006). *The network society: from knowledge to policy*. Center for Transatlantic Relations, Paul H. Nitze School of Advanced International Studies, Johns Hopkins University, Baltimore, MD.

- Chambers, E. (1728). *Cyclopædia, or, An universal dictionary of arts and sciences : containing the definitions of the terms, and accounts of the things signify'd thereby, in the several arts, both liberal and mechanical, and the several sciences, human and divine : the figures, kinds, properties, productions, preparations, and uses, of things natural and artificial : the rise, progress, and state of things ecclesiastical, civil, military, and commercial : with the several systems, sects, opinions, &c : among philosophers, divines, mathematicians, physicians, antiquaries, criticks, &c : the whole intended as a course of antient and modern learning*, volume 1, chapter Preface. Chambers, Ephraim, <http://digicoll.library.wisc.edu/cgi-bin/HistSciTech/HistSciTech-idx?type=article&did=HistSciTech.Cyclopaedia01.i0005&id=HistSciTech.Cyclopaedia01&isize=M>.
- Chisholm, H. (1910). Editorial introduction. In Chisholm, H. and Hooper, F. H., editors, *The Encyclopædia Britannica: A dictionary of arts, sciences, literature, and general information.*, volume 1, pages xi–xxiv. University Press, New York, Cambridge, England, eleventh edition.
- Clark, M., Ruthven, I., and O'Brian Holt, P. (2009). The evolution of genre in Wikipedia. *Journal for Language Technology and Computational Linguistics*, 24(1):1–22.
- Clark, R. (2001). A million pages and counting—recent German encyclopedia research. *Eighteenth-Century Studies*, 34(3):461–465.
- Collison, R. L. (1966). *Encyclopaedias: their history throughout the ages: a bibliographical guide with extensive historical notes to the general encyclopaedias issued throughout the world from 350 B.C. to the present day*. Hafner Publishing Company, New York, NY.
- Crawford, H. (2001). Encyclopedias. In R. Bopp, L. C. S., editor, *Reference and information services: an introduction*, pages 433–459. Libraries Unlimited, Englewood, CO, 3 edition.
- Crowston, K. and Williams, M. (2000). Reproduced and emergent genres of communication on the World Wide Web. *The Information Society*, 16:201–215.
- Deleuze, G. and Guattari, F. (1980 [1987]). *A Thousand Plateaus: Capitalism and Schizophrenia*. University of Minnesota Press, Minneapolis, MN. Trans. Brian Massumi.
- Doody, A. (2009). Pliny's natural history: Enkuklios paideia and the ancient encyclopedia. *Journal of the History of Ideas*, 70(1):1–21.
- Dooley, P. L. (2010). Wikipedia and the two-faced professoriate. In *WikiSym '10*, Gdańsk, Poland. ACM.
- Durbin, P. T. (1996). Encyclopedias and the integration of knowledge. *Social Epistemology*, 10(1):123–133.
- Economist, T. (2008). The battle for wikipedia's soul. <http://www.economist.com/node/10789354>.
- Einbender, H. (1964). *The Myth of the Britannica*. Grove Press, New York, NY.
- Emigh, W. and Herring, S. C. (2005). Collaborative authoring on the web: A genre analysis of online encyclopedias. *Hawaii International Conference on System Sciences*, 4:99a.

- Encyclopædia Britannica (1771). Preface. In Smellie, W., editor, *Encyclopædia Britannica; or, a dictionary of arts and sciences, compiled upon a new plan. In which the different sciences and arts are digested into distinct treatises or systems; and the various technical terms, &c. are explained as they occur in the order of the Alphabet.*, volume 1, pages v–vi. A. Bell & C. Macfarquhar, Edinburgh, first edition.
- Encyclopædia Britannica (1778). Preface. In Macfarquhar, C. and Tytler, J., editors, *Encyclopædia Britannica; or, a dictionary of arts and sciences, &c. On a plan entirely new: By which, the different sciences and arts are digested into distinct treatises or systems; comprehending the history, theory, and practice of each, according to the latest discoveries and improvements; and full explanations given of the various detached parts of knowledge, whether relating to natural and artificial objects, or to matters ecclesiastical, civil, military, commercial, &c. together with a description of all the countires, cities, principal mountains, seas, rivers, &c. throughout the world; A general history, ancient and modern, of the different empires, kingdoms, and states; and an account of the lives of the most eminent persons in every nation, from the earliest ages down to the present times.*, pages iii–viii. J. Balfour and Co., Edinburgh, second edition.
- Encyclopædia Britannica (1790). General preface. In Macfarquhar, C. and Gleig, G., editors, *Encyclopædia Britannica; or, a dictionary of arts, sciences, and miscellaneous literature, on a plan entirely new. By which the different sciences and arts are digested into the form of distinct treatises or systems, comprehending the history, theory, and practice, of each, according to the latest discoveries and improvements; and full explanations given of the various detached parts of knowledge, whether relating to natural and artificial objects, or to matters ecclesiastical, civil, military, commercial, &c. together with a description of all the countires, cities, principal mountains, seas, rivers &c. throughout the world. A general history, ancient and modern, of the different empires, kingdoms, and states; and an account of the lives of the most eminent persons in every nation, from the earliest ages down to the present times.*, volume 1, pages i–viii. James Moore, No. 45 College-Green, Dublin, third edition.
- Encyclopædia Britannica (1878). *Encyclopædia Britannica: A dictionary of arts, sciences, and general literature*, volume 1, pages v–viii. Thomas Spencer Baynes and W. Robertson Smith (Eds.) Samuel L. Hall, New York, ninth edition.
- Erickson, T. (2000). Making sense of computer-mediated communication (cmc): Conversations as genres, cmc systems as genre ecologies. In *Proceedings of the 33rd Annual Hawaii International Conference on System Sciences*, 2000.
- Fairclough, N. (1995). *Media Discourse*, chapter Intertextuality in the News. E. Arnold, London, UK; New York, NY.
- Fallis, D. (2009). Introduction: The epistemology of mass collaboration. *Episteme*, 6:1–7.
- Fallis, D. (2011). *Social epistemology: Essential readings*, chapter Wikipistemology, pages 297–313. Oxford University Press, Oxford, UK; New York, NY.
- Featherstone, M. and Venn, C. (2006). Problematizing global knowledge and the new encyclopaedia project: An introduction. *Theory, Culture & Society*, 23:1–20.

- Ferris, D. S. (2003). Post-modern interdisciplinarity: Kant, Diderot and the encyclopedic project. *MLN*, 118(5).
- Flanagin, A. J. and Metzger, M. J. (2011). From Encyclopædia Britannica to Wikipedia. *Information, Communication & Society*, 14(3):355–374.
- Flöck, F., Vrandečić, D., and Simperl, E. (2011). Towards a diversity-minded Wikipedia. In *WebSci'11*. ACM.
- Forte, A. and Bruckman, A. (2005). Why do people write for wikipedia? incentives to contribute to open-content publishing. In *GROUP '05: Proceedings of the 2005 International ACM SIGGROUP conference on Supporting group work*, pages 6–9.
- Foucault, M. (1997). *Ethics: Subjectivity and Truth*. New Press, New York, NY.
- Fowler, R. L. (1997). Definitions and theoretical problems. In *Pre-modern encyclopaedic texts: proceedings of the second COMERS Congress, Groningen, 1-4 July 1996*, pages 1–29. Brill, Leiden, NY.
- Frey, L. R., Botan, C. H., and Kreps, G. L. (2000). Textual analysis. In Eid, M. and Legace, M., editors, *Communication Research Methods: Quantitative and Qualitative Approaches*, pages 561–592. Pearson, Boston.
- Galloway, A. R. (2007). *The Exploit*. University of Minnesota Press, Minneapolis, MN.
- Genette, G. (1992). *The architext: an introduction*. University of California Press, Berkeley, CA.
- Genette, G. (1997). *Paratexts: thresholds of interpretation*. Cambridge University Press, Cambridge, UK; New York, NY.
- Genette, G. and Maclean, M. (1991). Introduction to the paratext. *New Literary History*, 22(2):pp. 261–272.
- Giles, J. (2005). Internet encyclopaedias go head to head. *Nature*, 438(15):990–901.
- Giltrow, J. and Stein, D. (2009). *Genres in the Internet: Issues in the Theory of Genre*, volume 188, chapter Genres in the Internet: Innovation, evolution, and genre theory, pages 1–26. John Benjamins Publishing Company.
- Glott, R., Schmidt, P., and Ghosh, R. (2010). *Analysis Of Wikipedia Survey Data*. Wikipediasurvey.org.
- Habermas, J. (1954 [1991]). *The structural transformation of the public sphere: An inquiry into a category of bourgeois society*. The MIT Press.
- Hadot, I., Treves, J. A., and Gage, J. C. (1997). Greek philosophy and encyclopedic knowledge. *Diogenes*, 45:33–47.
- Hall, S. (1980). *Culture, media, language*, chapter Encoding / Decoding, pages 123–148. Unwin Hyman.

- Heidegger, M. (1977). *The Question Concerning Technology, and Other Essays*. Harper & Row, New York, NY.
- Hutcheon, L. (1986). Postmodern paratextuality and history. *Texte - Revue de Critique et de Theorie Litteraire*, 5(6):301–312.
- Hutchins, R. M. (1984). Foreward. In Goetz, P. W. and Sutton, M., editors, *The New Encyclopædia Britannica: in 30 volumes*, volume 1, pages x–xi. Encyclopædia Britannica, Chicago, Auckland, Geneva, London, Manila, Paris, Rome, Seoul, Sydney, Tokyo, Toronto, fifteenth edition.
- Imbert, C., Davis, D. L., and Gage, J. C. (1997). Plato, the mirror of the world and the book. *Diogenes*, 45:7–22.
- Jackson, H. (2001). *Marginalia*. Yale University Press, New Haven.
- Jackson, S. L. (1977). Towards a history of the encyclopedia: From Amenemope of Egypt to the collapse of Greek in Rome. *The Journal of Library History (1974-1987)*, 12(4):342–358.
- Jacob, C., Treves, J. A., and Gage, J. C. (1997a). At the origins of the encyclopedic dream. *Diogenes*, 45:1–5.
- Jacob, C., Treves, J. A., and Gage, J. C. (1997b). The library and the book : Forms of Alexandrian encyclopedism. *Diogenes*, 45:63–82.
- Jamieson, K. (1975). Antecedent genre as rhetorical constraint. *Quarterly Journal of Speech*, 61(4):406–415.
- Jesus, R. (2010). What cognition does for wikis. In *WikiSym '10: Proceedings of the 6th International Symposium on Wikis and Open Collaboration*, pages 1–4, New York, NY. ACM.
- Jesus, R., Schwartz, M., and Lehmann, S. (2009). Bipartite networks of Wikipedia's articles and authors: a meso-level approach. In *WikiSym '09*.
- Kafker, F. (1973). The recruitment of the encyclopedists. *Eighteenth-Century Studies*, 6(4):452–461.
- Kafker, F. and Loveland, J. (2009). *The Early Britannica: the Growth of an Outstanding Encyclopedia*. Voltaire Foundation, Oxford, UK.
- Kamberelis, G. (1995). Genre as institutionally informed social practice. *Journal of Contemporary Legal Issues*, 6:115–171.
- Katz, B. (1998). *Cuneiform to Computer*. Scarecrow Press, Metuchen, NJ.
- Keen, A. (2007). *The Cult of the Amateur*. Currency Doubleday, New York, NY.
- Kennedy, K. (2009). *Textual curators and writing machines: Authorial agency in encyclopedias, print to digital*. PhD Dissertation, University of Minnesota, Minneapolis, MN.
- Kinser, S. (1990). *Rabelais's carnival: Text, context, metatext*. University of California Press, Berkeley, CA; Richmond, CA.

- Kittur, A., Chi, E., Pendleton, B., Suh, B., and Mytkowicz, T. (2007a). Power of the few vs. wisdom of the crowd: Wikipedia and the rise of the bourgeoisie. In *25th Annual ACM Conference on Human Factors in Computing Systems*, San Jose, CA.
- Kittur, A., Suh, B., Pendleton, B. A., and Chi, E. H. (2007b). He says, she says: Conflict and coordination in Wikipedia. In *CHI 2007*, San Jose, CA.
- Kogan, H. (1958). *The Great EB: The Story of the Encyclopaedia Britannica*. University of Chicago Press, Chicago, IL.
- Kools, M., Ruiter, R. A., van de Wiel, M. W., and Kok, G. (2008). The effects of headings in information mapping on search speed and evaluation of a brief health education text. *Journal of Information Science*, 34:833.
- Korzybski, A. (1958). *Science and Sanity: An Introduction to Non-Aristotelian Systems and General Semantics*. International Non-Aristotelian Library. International Non-Aristotelian Library Publishing Company.
- Kostakis, V. (2010). Identifying and understanding the problems of wikipedia's peer governance: The case of inclusionists versus deletionists. *First Monday*, 15(3).
- Kreeft, P. (1990). *A Summa of the Summa*. Ignatius Press, San Francisco, CA.
- Krippendorff, K. (2004). *Content Analysis: an introduction to its methodology*. Sage, Thousand Oaks, CA.
- Kwasnik, B. H. and Crowston, K. (2005). Genres of digital documents. *Information Technology & People*, 18(2):76–88.
- Laniado, D. and Tasso, R. (2011). Co-authorship 2.0: patterns of collaboration in wikipedia. In *Proceedings of the 22nd ACM conference on Hypertext and hypermedia*, HT '11, pages 201–210, New York, NY. ACM.
- Laniado, D., Tasso, R., Volkovich, Y., and Kaltenbrunner, A. (2011). When the Wikipedians talk: Network and tree structure of Wikipedia discussion pages. In *Fifth International AAAI Conference on Weblogs and Social Media*. Association for the Advancement of Artificial Intelligence.
- Lanier, J. (2010). *You Are Not a Gadget: a Manifesto*. Knopf, New York, NY.
- Latour, B. (1992 [2009]). *Technology and society: building our sociotechnical future*, chapter Where are the missing masses, sociology of a few mundane artefacts, pages 151–180. MIT Press, Cambridge, MA.
- Latour, B. (1999). *Pandora's hope: essays on the reality of science studies*, chapter A collective of humans and non-humans. Harvard University Press, Cambridge, MA.
- Latour, B. and Law, J. E. (1991). *A Sociology of Monsters. Essays on Power, Technology and Domination*, chapter Technology is society made durable, pages 103–130. Routledge, London, UK.
- Latour, B. and Venn, C. (2002). Morality and technology the end of the means. *Theory, Culture & Society*, 19(5-6):247–260.

- Lefevre, A. (1985). Systems in evolution: Historical relativism and the study of genre. *Poetics Today*, 6(4):pp. 665–679.
- Lefèvre, T., Jensen, C., and Korsgaard, T. (2009). *Trust Management III: IFIP Advances in Information and Communication Technology*, volume 300, chapter WRS: The Wikipedia Recommender System, pages 298–301. Springer Berlin Heidelberg.
- Leuf, B. and Cunningham, W. (2001). *The Wiki way: quick collaboration on the Web*. Addison-Wesley, Boston, MA.
- Levine, E. (2007). *Television: the critical view*, chapter Toward a paradigm for media production research: behind the scenes at General Hospital, pages 133–149. Oxford University Press, New York, NY.
- Lévy, P. (1994 [1997]). *Collective Intelligence: Mankind's Emerging World in Cyberspace*. Perseus Books, Cambridge, MA. Trans. Robert Bonono.
- Lévy, P. (2011). *The semantic sphere: computation, cognition and information economy*. ISTE Wiley, London, UK; Hoboken, NJ.
- Lih, A. (2009). *The Wikipedia revolution: how a bunch of nobodies created the world's greatest encyclopedia*. Aurum, London, UK.
- Maclean, M. (1991). Pretexts and paratexts: The art of the peripheral. *New Literary History*, 22(2):pp. 273–279.
- Manovich, L. (2000). Database as a genre of new media. *AI & Society*, 14:176–183. 10.1007/BF01205448.
- Martin, J. (1992). *English text: system and structure*. John Benjamins Pub. Co., Philadelphia, PA.
- Martin, J. and Rose, D. (2003). *Working with discourse: Meaning beyond the clause*. Continuum, London, UK; New York, NY.
- Marx, K. and Engels, F. (1846 [1970]). *The German Ideology*. International Publishers Co.
- McGann, J. J. (1991). *The textual condition*. Princeton University Press, Princeton, NJ.
- Miller, C. and Shepherd, D. (2009). Questions for genre theory from the blogosphere. *Pragmatics & beyond. New series*, 188:263–290.
- Miller, C. R. (1984). Genre as social action. *Quarterly Journal of Speech*, 70(2):151–167.
- Milne, D. (2010). A link-based visual search engine for Wikipedia. In *New Zealand Computer Science Research Student Conference*. University of Waikato.
- Mishler, E. (1984). *The discourse of medicine: Dialectics of medical interviews*. Ablex Publishing Corporation, Norwood, NJ.
- Morely, J. ([1905] 1886). *Diderot and the Encyclopaedists*, volume 1. Macmillan and Co., Limited, London, UK; New York, NY.

- Myers, K. S. (2006). Wikimmunity: Fitting the communications decency act to Wikipedia. *Harvard Journal of Law & Technology*, 20(1).
- Napier, M. (1842). Preface. In Napier, M. and Browne, J., editors, *The Encyclopedia Britannica: A dictionary of arts, sciences, and general literature. Seventh edition, with preliminary dissertations on the history of the sciences, and other extensive improvements and additions; including the late supplement, a general index, and numerous engravings.*, volume 1, pages v–xliii. Adam and Charles Black, Edinburgh, seventh edition.
- Nemoto, K., Gloor, P. A., and Laubacher, R. (2011). Social capital increases efficiency of collaboration among Wikipedia editors. In *22nd ACM conference on Hypertext and hypermedia*.
- Neuman, W. (2000). Qualitative and quantitative sampling. In Eid, M. and Legace, M., editors, *Communication Research Methods: Quantitative and Qualitative Approaches*, pages 345–370. Pearson, Toronto, ON.
- Noinville, J.-B. D. D. (1758). *Table alphabetique des dictionnaires, en toutes sortes de Langues & sur toutes sortes de Sciences & d'Arts*. Chez Hug. Chaubert [et] Herissant.
- Nordloh, D. J. (1998). General reference works. *American Literary Scholarship*, pages 517–528.
- Paling, S. (2002). Thresholds of access: Paratextuality and classification. *Journal of Education for Library and Information Science*, 43(2):pp. 134–143.
- Papert, S., Harel, I., and of Technology. Epistemology & Learning Research Group, M. I. (1991). *Constructionism: research reports and essays, 1985-1990*. Cognition and Computing Series. Ablex Publishing Corporation.
- Peeters, B. (2000). *The lexicon - encyclopedia interface: Current research in the semantics / pragmatics interface*, chapter Setting the scene. Recent milestones in the lexicon-encyclopedia debate, pages 1–52. Elsevier Science, Oxford.
- Pentzold, C. (2009). Fixing the floating gap: The online encyclopaedia Wikipedia as a global memory place. *Memory Studies*, 2:255–272.
- Pentzold, C. and Seidenglanz, S. (2006). Foucault@wiki: First steps towards a conceptual framework for the analysis of wiki discourses. In *WikiSym'06*.
- Preece, W. E. (1967). Editor's preface. In Preece, W. E., editor, *Encyclopedia Britannica: A new survey of universal knowledge*, volume 1, page vii. William Benton, Chicago, London, Toronto, Geneva, Sydney, Tokyo, Manila, fourteenth edition.
- Priedhorsky, R., Chen, J., Lam, S. T. K., Panciera, K., Terveen, L., and Riedl, J. (2007). Creating, destroying, and restoring value in wikipedia. In *GROUP'07*, pages 259–268, Sanibel Island, FL. ACM.
- Quintilian (77CE [1856, 2006]). Institutes of oratory. In L. Honeycutt; J.S. Watson, T., editor, <http://rhetoric.eserver.org/quintilian>. <http://rhetoric.eserver.org/quintilian>, 2012 edition.

- Rahman, M. M. (2006). *Essays analyzing blogs and Wikipedia*. PhD Dissertation, University of Kansas, Lawrence, KS.
- Rainie, L. and Tancer, B. (2007). 36% of online American adults consult Wikipedia. *Pew Internet & American Life Project Survey*.
- Reagle, J. (2007). Do as I do: Authorial leadership in Wikipedia. In *WikiSym'07*, pages 143–155. ACM.
- Reagle, J. (2008). *In good faith: Wikipedia collaboration and the pursuit of the universal encyclopedia*. PhD Dissertation, New York University, New York, NY.
- Ribémont, B. (1997). On the definition of the encyclopaedic genre in the middle ages. In *Pre-modern encyclopaedic texts: proceedings of the second COMERS Congress, Groningen, 1-4 July 1996*, pages 47–61. Brill, Leiden, NY.
- Rosenzweig, R. (2006). Can history be open source? Wikipedia and the future of the past. *The Journal of American History*, 93(1):117–146.
- Ruth, A. and Houghton, L. (2009). The wiki way of learning. *Australasian Journal of Educational Technology*, 25(2):135–152.
- Sanger, L. (2009). The fate of expertise after wikipedia. *Episteme*, 6(01):52–73.
- Santini, M., Mehler, A., and Sharoff, S. (2011). *Genres on the Web*, volume 42 of *Text, Speech and Language Technology*, chapter Riding the Rough Waves of Genre on the Web, pages 3–30. Springer Netherlands.
- Shirky, C. (2008). *Here comes everybody : the power of organizing without organizations*. Penguin Press, New York, NY.
- Shirky, C. (2010). *Cognitive surplus : creativity and generosity in a connected age*. Penguin Press, New York, NY.
- Solon, O. (2011). Q&A: Jimmy Wales reflects on a decade of Wikipedia. *Wired*.
- Spiekermann, E. and Ginger, E. (2003). *Stop stealing sheep and find out how type works*. Adobe Press, Berkeley, CA.
- Stecchini, L. C. (1962). On encyclopedias in time and space. *American Behavioral Scientist*, 6:3–6.
- Steup, M. (2005 [2010]). Epistemology. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. <http://plato.stanford.edu/archives/spr2010/entries/epistemology/>, spring 2010 edition.
- Stewart, D. (1853). Preface, containing some critical remarks on the discourse prefixed to the French “encyclopédie”. In Traill, T. S., editor, *The Encyclopedia Britannica: A dictionary of arts, sciences, and general literature. Eighth edition, with extensive improvements and additions; and numerous engravings.*, volume 1, pages 1–12. Adam and Charles Black, Edinburgh, eighth edition.

- Stover, C. F. (1962). Change and rationality in encyclopedism. *American Behavioral Scientist*, 6:35–38.
- Stvilia, B., Twidale, M. B., Gasser, L., and Smith, L. C. (2005). Information quality discussions in Wikipedia. In *International Conference on Knowledge Management*.
- Sullivan, L. E. (1990). Circumscribing knowledge: Encyclopedias in historical perspective. *The Journal of Religion*, 70(3):315–339.
- Tapscott, D. (2006). *Wikinomics: How Mass Collaboration Changes Everything*. Portfolio, New York, NY.
- Thom-Santelli, J., Cosley, D., and Gay, G. (2009). What’s mine is mine: Territoriality in collaborative authoring. In *CHI 2009 — Personal and Online Information*, Boston, MA.
- Thomas, G. (1992). *A position to command respect: women and the Eleventh Britannica*. Scarecrow Press, Metuchen, NJ.
- Tschichold, J. (1928 [2006]). *The New Typography*. University of California Press, London, UK; Berkeley, CA.
- Ulrike Pfeil, Panayiotis Zaphiris, C. S. A. (2006). Cultural differences in collaborative authoring of Wikipedia. *Journal of Computer-Mediated Communication*, 12(1):<http://jcmc.indiana.edu/vol12/issue1/pfeil.html>.
- University of Chicago Press (1906). *Manual of Style: Being a compilation of the typographical rules in force at the University of Chicago Press*. University of Chicago Press, Chicago, IL.
- University of Chicago Press (1982). *The Chicago Manual of Style*. University of Chicago Press, Chicago, IL.
- van Doren, C. (1962). The idea of an encyclopedia. *American Behavioral Scientist*, 6:23–26.
- Van Leeuwen, T. (1987). Generic strategies in press journalism. *Australian Review of Applied Linguistics*, 10(2):199–220.
- van Loon, J. (2000). Organizational spaces and networks. *Space and Culture*, 2(4-5):109–112.
- van Loon, J. (2006). Network. *Theory, Culture & Society*, 23:307–314.
- Viégas, F. B., Wattenberg, M., Kriss, J., and van Ham, F. (2007a). Talk before you type: Coordination in Wikipedia. In *HICSS-40*, pages 1–10, New York, NY.
- Viégas, F. B., Wattenberg, M., and Kushal, D. (2004). Studying cooperation and conflict between authors with history flow visualizations. In *CHI '04: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 575–582, New York, NY. ACM.
- Viégas, F. B., Wattenberg, M., and McKeon, M. M. (2007b). The hidden order of Wikipedia. In Schuler, D., editor, *Online Communities and Social Computing*, pages 445–454. HCII 2007, Springer-Verlag Berlin Heidelberg.

- Washingtonpost.com (Retrieved April 25, 2013). Ap sources: Boston bomb suspects from russia region near chechnya, lived in us at least 1 year. http://articles.washingtonpost.com/2013-04-19/politics/38658080_1_ap-sources-boston-bomb-chechnya.
- Wellman, B., Quan-Haase, A., Boase, J., Chen, W., Hampton, K., de Diaz, I. I., and Miyata, K. (2003). The social affordances of the internet for networked individualism. *Journal of Computer Mediated Communication*, 8(3).
- Wenger, E. (1999). *Communities of Practice: Learning, Meaning, and Identity*. Cambridge University Press, Cambridge, UK; New York, NY.
- Westerman, W. (2009). *Folklore and the Internet: Vernacular expression in a digital world*, volume 35, chapter Epistemology, the Sociology of Knowledge, and the Wikipedia Userbox Controversy, pages 123–158. All USU Press Publications.
- Wikimedia Foundation, I. (2011). Financial statements June 30, 2011 and 2010. http://wikimediafoundation.org/wiki/Finance_report.
- Wikipedia (2001a [2011]c). Wikipedia:Encyclopedia. <http://en.wikipedia.org/w/index.php?title=Encyclopedia&oldid=441672456>.
- Wikipedia (2001b [2011]e). Wikipedia:Gazetteer. <http://en.wikipedia.org/w/index.php?title=Gazetteer&oldid=439295201>.
- Wikipedia (2002 [2011]f). Wikipedia:Main Page. http://en.wikipedia.org/w/index.php?title=Main_Page&oldid=447996010.
- Wikipedia (2003 [2011]b). Wikipedia:almanac. <http://en.wikipedia.org/w/index.php?title=Almanac&oldid=440915328>.
- Wikipedia (2004a [2012]d). Wikipedia:Summary style. http://en.wikipedia.org/w/index.php?title=Wikipedia:Summary_style&oldid=524300750.
- Wikipedia (2004b [2011]a). List of Wikipedias. http://meta.wikimedia.org/w/index.php?title=List_of_Wikipedias&oldid=3029166.
- Wikipedia (2004c [2012]b). Wikipedia:Categorization. <http://en.wikipedia.org/w/index.php?title=Wikipedia:Categorization&oldid=475996526>.
- Wikipedia (2005a [2011]d). Wikipedia:five pillars. http://en.wikipedia.org/w/index.php?title=Wikipedia:Five_pillars&oldid=443049062.
- Wikipedia (2005b [2010]). Wikipedia:Classifications. <http://en.wikipedia.org/w/index.php?title=Wikipedia:Classification&oldid=401303004>.
- Wikipedia (2006 [2012]c). Wikipedia:Overcategorization. <http://en.wikipedia.org/w/index.php?title=Wikipedia:Overcategorization&oldid=524766281>.
- Wikipedia (2007a [2012]e). WikiProject:Outlines. http://en.wikipedia.org/w/index.php?title=Wikipedia:WikiProject_Outlines&oldid=476236260.
- Wikipedia (2007b [2012]a). Category:contents. <http://en.wikipedia.org/w/index.php?title=Category:Contents&oldid=476766537>.

- Wikipedia (2007c [2009]). Wikipedia:purpose. <http://en.wikipedia.org/wiki/Wikipedia:Purpose>.
- Wikipedia (2012a). Special:Statistics. <http://en.wikipedia.org/wiki/Special:Statistics>.
- Wikipedia (2012b). Wikipedia:wikipedians demographics. <http://en.wikipedia.org/wiki/Wikipedia:Wikipedians#Demographics>.
- Wilkinson, D. M. and Huberman, B. A. (2008). *Cooperation and Quality in Wikipedia*. PhD Dissertation, HP Labs, Palo Alto, CA.
- Williams, T. and Spyridakis, J. (1992). Visual discriminability of headings in text. *Professional Communication, IEEE Transactions on*, 35(2):64–70.
- Withers, C. W. J. (1996). Encyclopaedism, modernism and the classification of geographical knowledge. *Transactions of the Institute of British Geographers*, 21(1):275–298.
- Wolff, G. (1976). Britannica 3, failures of. *The Atlantic Monthly*, pages 107–110.
- Wright, A. (2008). *Glut: mastering information through the ages*. Cornell University Press, Ithaca, NY.
- Wu, G., Harrigan, M., and Cunningham, P. (2011). A characterization of wikipedia content based on motifs in the edit graph. Technical report, University College Dublin, Dublin, Ireland.
- Yeo, R. (1991). Reading encyclopedias: Science and the organization of knowledge in British dictionaries of arts and sciences, 1730–1850. *Isis*, 82(1):24–49.
- Yeo, R. (1996). Ephraim chambers’s cyclopædia (1728) and the tradition of commonplaces. *Journal of the History of Ideas*, 57(1):157–175.
- Yeo, R. (2003). A solution to the multitude of books: Ephraim Chambers’s “Cyclopaedia” (1728) as “the best book in the universe”. *Journal of the History of Ideas*, 64(1):61–72.
- Yeo, R. (2007). Lost encyclopedias: Before and after the Enlightenment. *Book History*, 10:47–68.
- Youtube (2012). Ward cunningham, wiki founder, speaks at hacking social impact unconference, youtube. <http://www.youtube.com/watch?v=Gnmf2DLPIwU>.
- Zhu, H., Kraut, R. E., Wang, Y.-C., and Kittur, A. (2011). Identifying shared leadership in Wikipedia. In *Proceedings of the 2011 annual conference on Human factors in computing systems*, CHI ’11, pages 3431–3434, New York, NY. ACM.

12 Appendix

Figure 1
Chambers' *Cyclopedia*, 1728

Table of Arrangements

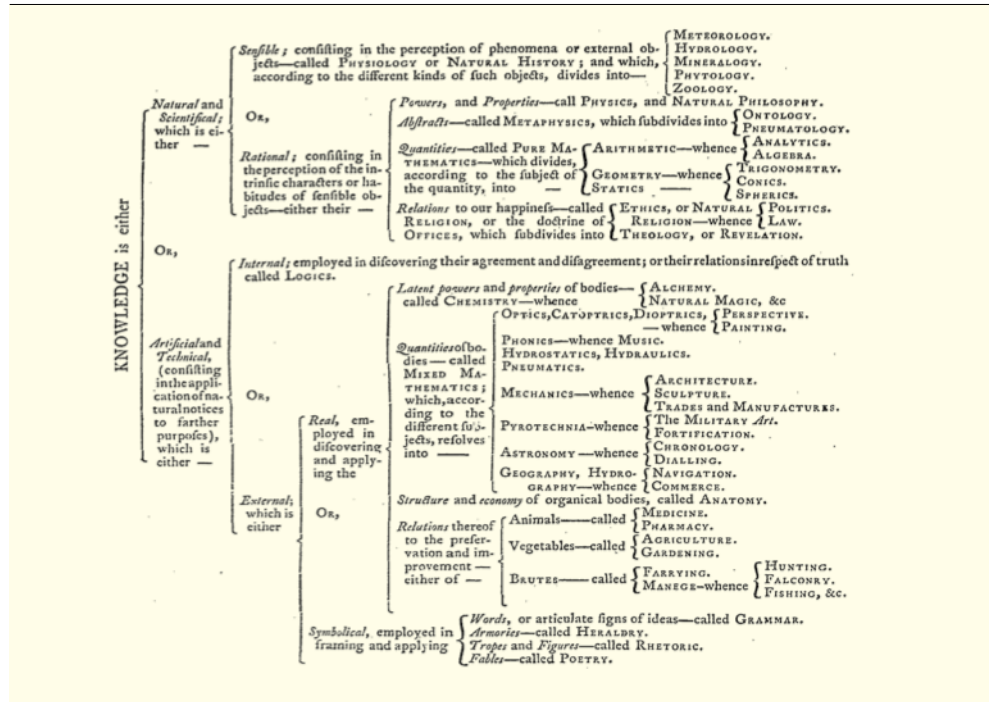


Figure 2
Diderot's *Encyclopédie*, 1751

Système figuré des connaissances humaines (The System of Human Knowledge)

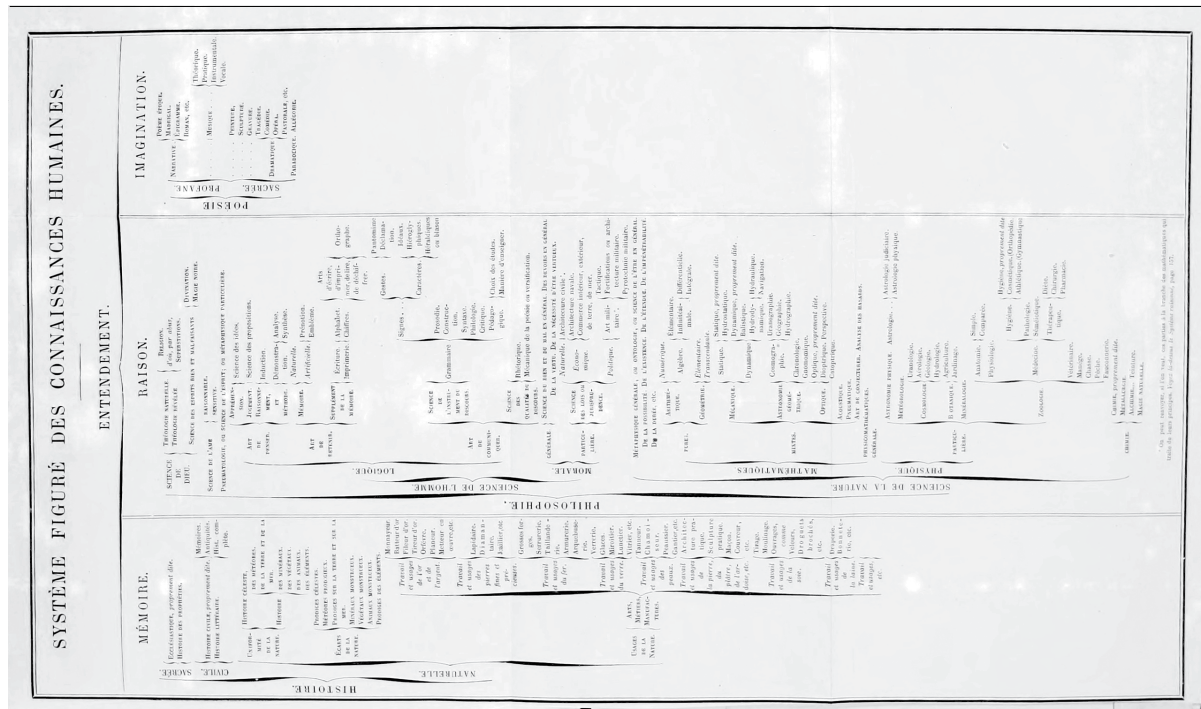




Table 1

Ranking of top 300 edited Wikipedia pages

Rank	Edits	Article Name	Rank	Edits	Article Name
1	764923	Wikipedia:Administrator intervention against vandalism	49	35248	Wikipedia:Pages needing translation into English
2	514289	Wikipedia:Administrators' noticeboard/Incidents	50	34790	Wikipedia:Department of Fun/Word Association
3	343544	Wikipedia:Sandbox	51	34674	User talk:OrphanBot
4	326338	Wikipedia:Sandbox/Archive	52	33899	Wikipedia:Sandbox/Archive2
5	316017	User:Cyde/List of candidates for speedy deletion/Subpage	53	32726	Wikipedia:Wikiquette alerts
6	254049	Template talk:Did you know	54	32630	Wikipedia talk:WikiProject Football
7	252598	Wikipedia:Reference desk/Science	55	32583	List of World Wrestling Entertainment personnel
8	233497	Wikipedia:WikiProject Spam/LinkReports	56	32152	Wikipedia:Copyright problems
9	212135	Wikipedia:Reference desk/Miscellaneous	57	32102	Wikipedia:Arbitration/Requests/Enforcement
10	201699	Wikipedia:Requests for page protection	58	32052	Wikipedia:Sandbox/History
11	198495	Wikipedia:Help desk	59	31915	Wikipedia talk:WikiProject Professional wrestling
12	185419	Wikipedia:Reference desk/Humanities	60	31809	Wikipedia talk:Manual of Style
13	173607	Wikipedia:Administrators' noticeboard	61	31459	Wikipedia:Main Page/Errors
14	145615	Wikipedia:Introduction	62	31208	User:VeblenBot/SPERTable
15	127417	Wikipedia:Usernames for administrator attention	63	30878	Wikipedia talk:Manual of Style (dates and numbers)
16	105276	Wikipedia:Reference desk/Computing	64	30324	Wikipedia:Media copyright questions
17	101135	Wikipedia:Reference desk/Language	65	30262	Wikipedia
18	100170	Talk:Main Page	66	30206	Wikipedia:Tutorial/Formatting/sandbox
19	82317	Wikipedia:Usernames for administrator attention/Bot	67	30076	Wikipedia:Cleanup
20	82304	Wikipedia:Tutorial/Editing/sandbox	68	30003	Wikipedia talk:WikiProject Video games
21	82177	User:Cyde/List of current proposed deletions	69	28505	United States
22	79042	Wikipedia:Open proxy detection	70	28484	Wikipedia:Village pump archive 2004-09-26
23	78050	Wikipedia:Administrator intervention against vandalism/TB2	71	28457	Wikipedia:Editor assistance/Requests
24	77995	User:COIBot/LinkReports	72	28390	Wikipedia talk:Reference desk
25	75174	Wikipedia:Administrators' noticeboard/Edit warring	73	28386	Wikipedia:Deletion review
26	71775	Wikipedia:Village pump (technical)	74	28318	Talk:Global warming
27	71467	Wikipedia:Good article nominations	75	28306	Portal:Current events/Sports
28	70776	User:Cyde/List of candidates for speedy deletion	76	27184	Wikipedia:Village pump (miscellaneous)
29	69661	Wikipedia:Suspected copyright violations	77	27074	Wikipedia:Requests for investigation
30	69385	Wikipedia:Requested moves	78	27057	Talk:Sarah Palin
31	67390	Wikipedia:Reference desk/Mathematics	79	26973	User:ClueBot/PossibleVandalism
32	66510	Wikipedia:Village pump (policy)	80	26099	Wikipedia:Main Page/Protection
33	62878	Wikipedia:Changing username/Simple	81	25391	Michael Jackson
34	61972	User:VeblenBot/PERTable	82	25101	Wikipedia:New contributors' help page
35	56174	Wikipedia:Arbitration/Requests	83	25011	Wikipedia:Requests for adminship
36	55464	Wikipedia:Village pump (proposals)	84	24962	User:X!/RfX Report
37	54432	Wikipedia talk:Requests for adminship	85	24887	Wikipedia:Templates for discussion
38	52884	Wikipedia:In the news/Candidates	86	24634	User talk:SandyGeorgia
39	51142	User:Cyde/List of requests for unblock	87	24589	Wikipedia talk:WikiProject Military history
40	46456	User talk:Jimbo Wales	88	24505	Wikipedia:Featured article candidates
41	46299	Wikipedia:Reliable sources/Noticeboard	89	24454	Template:Toolserver
42	44169	George W. Bush	90	24373	Wikipedia talk:Criteria for speedy deletion
43	43736	Template:Did you know/Preparation area 1	91	23811	Wikipedia:Requested articles/Culture and fine arts
44	41500	Wikipedia:Biographies of living persons/Noticeboard	92	23762	Talk:Gaza War
45	40510	Wikipedia:Reference desk/Entertainment	93	23239	Wikipedia talk:No original research
46	40245	Wikipedia:Votes for deletion archive May 2004	94	23031	User:Tangotango/RfA Analysis/Report
47	39460	Wikipedia:Huggle/Whitelist	95	22974	Wikipedia:Bureaucrats' noticeboard/RfA Report
48	36299	Talk:Barack Obama	96	22795	Wikipedia talk:Non-free content
			97	22766	Wikipedia:Categories for discussion/Working

Table 1 (continued)

Ranking of top 300 edited Wikipedia pages

Rank	Edits	Article Name	Rank	Edits	Article Name
98	22749	Jesus	150	17370	Wikipedia:Redirects for discussion
99	22612	Wikipedia:Good articles	151	17252	Jehovah's Witnesses
100	22479	Wikipedia:Did you know/DYK hook count	152	17201	Real Madrid C.F.
101	22182	Talk:Intelligent design	153	17014	Wikipedia:Files for deletion
102	22088	Wikipedia:Changing username/Usurpations	154	16937	User:Cyde/List of old proposed deletions
103	22013	Wikipedia:Conflict of interest/Notice-board	155	16882	Wikipedia:Motto of the day/Nominations/In review
104	21911	Talk:George W. Bush	156	16840	User talk:Moonriddengirl
105	21665	Wikipedia:Articles for deletion/Old	157	16813	Islam
106	21334	Britney Spears	158	16715	Wikipedia:Bot requests
107	21300	World War II	159	16714	Talk:Evolution
108	21287	Wikipedia talk:Did you know	160	16710	Talk:Homeopathy
109	21007	Wikipedia talk:WikiProject Spam	161	16704	European Union
110	21002	Template:In the news	162	16608	User talk:Xeno
111	20882	Deaths in 2009	163	16498	FC Barcelona
112	20844	Adolf Hitler	164	16487	Muhammad
113	20821	Wii	165	16484	Wikipedia:Tutorial/Wikipedia links/sandbox
114	20598	User talk:Malleus Fatuorum	166	16450	Hurricane Katrina
115	20548	Catholic Church	167	16445	Christianity
116	20512	Barack Obama	168	16401	Talk:Muhammad
117	20458	Deaths in 2010	169	16390	User:SQL/Blanked
118	20456	Wikipedia:Historical archive/Sandbox	170	16371	Wikipedia talk:WikiProject Anime and manga
119	20291	Wikipedia:Lists of protected pages	171	16325	Elvis Presley
120	20064	Talk:Jesus	172	16309	User talk:YellowMonkey
121	20063	List of Total Drama series characters	173	16277	Wikipedia:Requested moves/current
122	19936	Wikipedia talk:Featured article candidates	174	16260	List of Omnitrix aliens
123	19781	Wikipedia:Article Collaboration and Improvement Drive	175	16178	User talk:DGG
124	19578	Talk:Climatic Research Unit email controversy	176	16145	Canada
125	19527	Wikipedia talk:WikiProject Mathematics	177	16075	Wikipedia:Categories for discussion
126	19389	Wikipedia:Graphic Lab/Illustration workshop	178	16026	User talk:Durova
127	19293	RuneScape	179	16007	Wikipedia:Possibly unfree files
128	19199	Wikipedia talk:Verifiability	180	15977	Wikipedia talk:Notability
129	19068	Wikipedia:Categories for discussion/Speedy	181	15834	List of programs broadcast by ABS-CBN
130	19047	Deaths in 2008	182	15817	User talk:AnmaFinotera
131	19041	2006 Lebanon War	183	15816	User talk:Alison
132	19003	PlayStation 3	184	15813	Led Zeppelin
133	18966	Wikipedia:BAG/Status	185	15741	Beyoncé Knowles
134	18923	User talk:Dr. Blofeld	186	15708	Xbox 360
135	18592	The Beatles	187	15685	Talk:British Isles
136	18555	Talk:Race and intelligence	188	15649	User talk:William M. Connolley
137	18480	Talk:September 11 attacks	189	15639	Mariah Carey
138	18416	Talk:Prem Rawat	190	15632	Anarchism
139	18350	Wikipedia:Featured picture candidates	191	15624	Wikipedia:Today's featured article/requests
140	18256	Template:Did you know	192	15596	Wikipedia talk:Biographies of living persons
141	18201	Deaths in 2007	193	15595	User talk:Raul654
142	18149	Global warming	194	15512	Eminem
143	18076	Wikipedia:Categories for discussion/User	195	15440	Ronald Reagan
144	18060	Talk:Catholic Church	196	15421	Roger Federer
145	17939	United Kingdom	197	15381	Wikipedia talk:WikiProject Comics
146	17890	Wikipedia:Requests for comment/User names	198	15299	Template:X1
147	17792	India	199	15283	United States presidential election, 2008
148	17735	Talk:Anarchism	200	15271	User talk:Rich Farmbrough
149	17588	User:Cyde/List of requests for help	201	15264	The Undertaker
			202	15203	New York City
			203	15195	Wikipedia talk:WikiProject Ice Hockey

Table 1 (continued)

Ranking of top 300 edited Wikipedia pages

Rank	Edits	Article Name	Rank	Edits	Article Name
204	15183	User talk:Gwen Gale	251	13712	User:ST47/AfD
205	15151	Wikipedia:Requested articles/Social sciences	252	13676	South Korea
206	15143	Blackout (Britney Spears album)	253	13665	Chicago
207	15101	September 11 attacks	254	13628	Heroes (TV series)
208	15093	Super Smash Bros. Brawl	255	13619	Sarah Palin
209	15091	Talk:International recognition of Kosovo	256	13614	Lost (TV series)
210	14974	Celtic F.C.	257	13535	Wikipedia talk:What Wikipedia is not
211	14870	London	258	13451	Wikipedia talk:WikiProject Cricket
212	14862	Israel	259	13413	Talk:2008 South Ossetia war
213	14800	Wikipedia talk:WikiProject Birds	260	13380	Britney Spears discography
214	14772	John Cena	261	13330	Russia
215	14768	Philippines	262	13320	Buddhism
216	14763	Wikipedia talk:Notability (fiction)	263	13275	Avatar: The Last Airbender
217	14734	Talk:Atheism	264	13215	iPod
218	14701	Talk:Chiropractic	265	13205	Harry Potter and the Deathly Hallows
219	14689	Red Hot Chili Peppers	266	13194	Talk:Transcendental Meditation
220	14669	Iran	267	13169	Wikipedia:Requested articles/music
221	14636	User talk:Bishonen	268	13160	2007
222	14635	Japan	269	13157	List of best-selling music artists
223	14545	Germany	270	13111	User talk:J.delanoy
224	14469	Wikipedia:Coordination/Requested moves	271	13078	Template:2009-2010 flu pandemic table
225	14432	World Wrestling Entertainment	272	13058	Wikipedia:Miscellany for deletion
226	14340	American Idol	273	13054	List of Total Nonstop Action Wrestling employees
227	14311	Talk:Michael Jackson	274	13036	Chelsea F.C.
228	14282	User talk:Tony Sidaway	275	13000	Kane (wrestler)
229	14246	Wikipedia:Requested moves/current-oldstyle	276	12973	John Kerry
230	14224	Wikipedia:Village pump (assistance)	277	12950	Edge (wrestler)
231	14207	Madonna (entertainer)	278	12947	2006 FIFA World Cup
232	14198	Talk:United States	279	12934	Wikipedia:Requested articles/Sports
233	14187	Scientology	280	12896	Harry Potter
234	14060	Wikipedia:Tutorial/Citing sources/sandbox	281	12895	Evolution
235	14054	Deaths in 2006	282	12885	Doctor Who
236	14034	User:Dragons flight/RFA summary	283	12860	User talk:Iridescent
237	13997	Iraq War	284	12859	Ireland
238	13958	Wikipedia:Fringe theories/Noticeboard	285	12834	Pakistan
239	13952	Gaza War	286	12817	Wikipedia:Requests for checkuser
240	13929	Akatsuki (Naruto)	287	12815	Triple H
241	13908	Bill Clinton	288	12745	Christina Aguilera
242	13908	User talk:Ryulong	289	12725	Pope John Paul II
243	13877	Sonic the Hedgehog (character)	290	12715	2008
244	13840	Guns N' Roses	291	12671	Australia
245	13834	Manchester United F.C.	292	12632	Neighbours
246	13819	Albert Einstein	293	12612	Wikipedia talk:Sandbox/Archive
247	13813	Talk:United States and state terrorism	294	12603	Paul McCartney
248	13790	Cannabis (drug)	295	12586	24 (TV series)
249	13746	World War I	296	12577	User talk:Dbachmann
250	13722	Organization XIII	297	12513	Pink Floyd
			298	12509	Scotland
			299	12452	Metallica
			300	12449	Brazil

Table 2

Ranking of top 24 edited Wikipedia pages that are also in the 1771 Encyclopedia Britannica

ArticleName*	Edits	Edits/600	Edits/24	Traffic2008	Traffic2009	Traffic2010	(2008+09+10)	Traffic	Edits+Traffic	DateCreated
United_States	28505	69	1	13073579	15017131	34180907	62271617	1	2	09/06/2001
India	17792	147	2	6840315	8196595	13500520	28537430	2	4	10/26/2001
Canada	16145	176	3	6055422	7018492	12988352	26062266	3	6	10/31/2001
Japan	14635	222	7	5901229	6708372	10299571	22909172	4	11	10/31/2001
Germany	14545	223	8	5227257	5865657	11556740	22649654	5	13	11/09/2001
London	14870	211	4	5442567	4334541	7208939	16986047	9	13	10/08/2001
Russia	13330	261	9	5233871	4995366	8579179	18808416	8	17	10/31/2001
Philippines	14768	215	5	3277688	4412550	5927867	13618105	13	18	05/26/2001
England	12435	301	12	4061586	4246921	10651218	18959725	7	19	10/17/2001
Brazil	12449	300	11	3842906	4503856	7462185	15808947	11	22	09/30/2001
Iran	14669	220	6	2505019	3970766	4518572	10994357	16	22	10/22/2001
France	11349	376	17	4666146	5055636	10738481	20460263	6	23	11/04/2001
Italy	11893	333	15	3995025	4267637	8428879	16691541	10	25	11/12/2001
Turkey	12094	316	14	3060920	3650223	5148004	11859147	14	28	09/02/2001
Scotland	12509	298	10	2452219	2661872	4820844	9934935	19	29	10/01/2001
Greece	12318	306	13	2685551	2823202	4542236	10050989	18	31	09/10/2001
Spain	10733	429	21	3789370	4046775	7599624	15435769	12	33	10/03/2001
Cat	11891	334	16	2976935	3147206	3296793	9420934	20	36	11/09/2001
Sweden	10653	434	23	2676544	3365725	5778178	11820447	15	38	10/01/2001
People's Republic_of_China	10686	433	22	2648801	3004113	5181818	10834732	17	39	10/23/2001
Portugal	11114	393	19	1739241	2000569	3519767	7259577	21	40	03/15/2001
Atheism	11229	385	18	1238113	1816361	2199648	5254122	25	43	11/13/2001
Democracy	10745	424	20	1776225	1807356	2336219	5919800	24	44	10/14/2001
Human	9296	595	25	1922493	2380383	2710486	7013362	22	47	10/03/2001
Dominican _Republic	10625	436	24	1911132	2131922	2860801	6903855	23	47	09/05/2001

* Some article names differ in the first edition of the Britannica: America, Turkey, Feline, China Dominica

Table 3

Ranking of 10 edited Wikipedia pages that are also in the 1771 Encyclopedia Britannica

ArticleName*	Edits	Edits/600	Edits/24	Traffic2008	Traffic2009	Traffic2010	(2008+09+10)	Traffic	Edits+Traffic	DateCreated
United_States	28505	69	1	13073579	15017131	34180907	62271617	1	2	09/06/2001
Canada	16145	176	3	6055422	7018492	12988352	26062266	3	6	10/31/2001
London	14870	211	4	5442567	4334541	7208939	16986047	9	13	10/08/2001
Cat	11891	334	16	2976935	3147206	3296793	9420934	20	36	11/09/2001
Atheism	11229	385	18	1238113	1816361	2199648	5254122	25	43	11/13/2001
Democracy	10745	424	20	1776225	1807356	2336219	5919800	24	44	10/14/2001
Human	9296	595	25	1922493	2380383	2710486	7013362	22	47	10/03/2001
Ocean	3926	N/A	26	885130	930307	1099086	2914523	26	52	06/18/2001
Black	3247	N/A	27	862127	932602	833177	2627906	27	54	08/13/2001
Cannon	2925	N/A	28	419543	365761	416022	1201326	28	56	11/09/2001

Table 4

Encyclopedia Britannica: Publication data

Edition	Volumes	Date	Editors
First	3	1768–1771	William Smellie
Second	10	1777–1784	Colin Macfarquhar, James Tytler
Third	18	1788–1797	Colin Macfarquhar, George Gleig
Supplement	2	1801	
Fourth	18+2	1801–1809	James Millar
Fifth	20	1815	James Millar
Sixth	20	1823–1824	Charles Maclaren
Supplement	1	1815–1824	
Seventh	21	1830–1842	Macvey Napier, James Browne
Eighth	22	1852–1860	Thomas Stewart Traill
Ninth	25	1875–1889	Thomas Spencer Baynes, W. Robertson Smith
Tenth	25+11	1902–1903	Hugh Chisholm, Sir Donald Mackenzie Wallace, Arthur T. Hadley, Franklin Henry Hooper
Eleventh	29	1911	Hugh Chisholm, Franklin Henry Hooper
Twelfth	29+3	1922	
Thirteenth	29+3+3	1926	
Fourteenth	23	1929–1973	J. L. Garvin, Franklin Henry Hooper 1929, 1930, 1932, 1933, 1936, 1937, 1938, 1939, 1940, 1941, 1942, 1943, 1944, 1945, 1946, 1947, 1948, 1949, 1950, 1951, 1952, 1953, 1954, 1955, 1956, 1957, 1958, 1959, 1960, 1961, 1962, 1963, 1964, 1965, 1966, 1967, 1968, 1969, 1970, 1971, 1972, 1973
Fifteenth	30	1974–1984	Board of Editors 1974, 1975, 1976, 1977, 1978, 1979, 1980, 1981, 1982, 1983, 1984
Fifteenth		1985–2010	Philip W. Goetz 1985, 1986, 1987, 1988, 1989, 1990, 1991, 1992, 1995, 1998, 2002, 2003, 2005, 2006, 2007, 2010
CD/DVD-ROM	N/A	1994, 1999, 2002, 2003, 2009, 2010, 2011	
Website	N/A	1994–	

*Dates prior to 1984 found in Propaedia, Outline of knowledge, Guide to the Britannica, 1984

Table 5

Publication dates of the Encyclopedia Britannica and the publications used for data collection

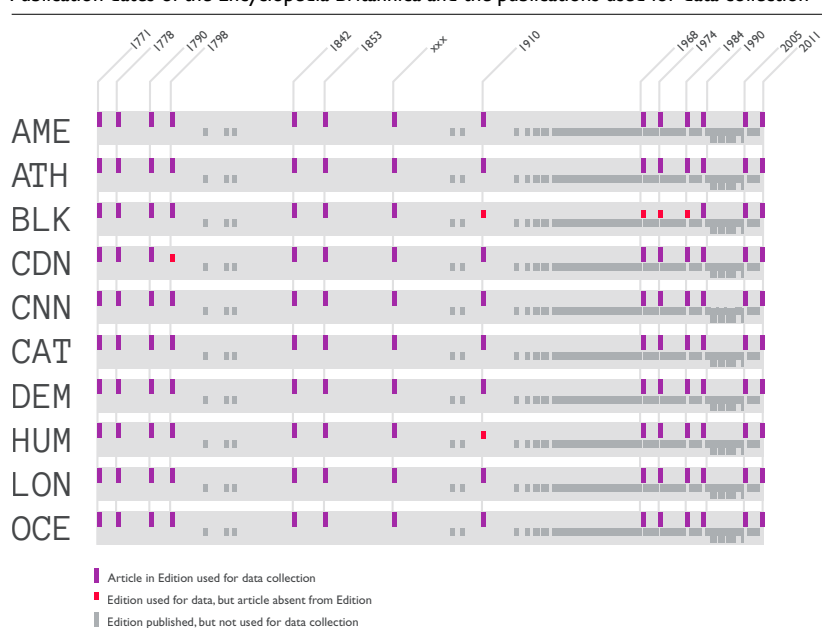


Table 6

Encyclopedia Britannica article corpus with bibliographic information

Article	Year	Edition	Volume	Pages	Actual Heading
America	1771	1	1	134	America
America	1778	2	1	288-308	America
America	1790	3	1	537-618	America
America	1798	3	1	538-618	America
America	1842	7	21	450-481	United States of North America
America	1853	8	21	419-442	United States of North America
America	1878	9	23	729-830	United States
America	1910	11	27	612-735	United States, The
America	1968	14	22	578-743	United States (of America)
America	1974	15	18	905-1003	United States (of America)
America	1984	15	18	905-1002	United States (of America)
America	1990	15	29	149-451	United States of America
America	2005	15	29	149-458	United States of America
America	2011	Web	-	EBchecked/topic/616563/United-States	United States of America
Atheism	1771	1	1	501	Atheist
Atheism	1778	2	2	854	Atheism/Atheist
Atheism	1790	3	2	604	Atheism/Atheist
Atheism	1798	3	2	604	Atheism/Atheist
Atheism	1842	7	4	113	Atheist
Atheism	1853	8	4	158-158	Atheist
Atheism	1878	9	23	234	[No Article] (via: Theism > Meanings of the word theism > Its generic sense > Its negatives)
Atheism	1910	11	2	827-828	Atheism
Atheism	1968	14	2	667-68	Atheism
Atheism	1974	15	2	258-262	Atheism
Atheism	1984	15	2	258-262	Atheism
Atheism	1990	15	26	611-614	Atheism (via: Systems of Religious and Spiritual Belief)
Atheism	2005	15	26	572-577 Belief, Systems of > Agnosticism)	Atheism (via: Religious and Spiritual
Atheism	2011	Web	-	EBchecked/topic/40634/atheism	atheism
Black	1771	1	1	558-559	Black
Black	1778	2	2	1169-1171	Black
Black	1790	3	3	253-255	Black
Black	1798	3	3	253-255	Black
Black	1842	7	4	665-658	Black
Black	1853	8	4	737-740	Black
Black	1878	9	7	579-580	Black Colours (via: Dyeing > Compound Colours)
Black	1910	11	-	-	[No article or Index item]
Black	1968	14	-	-	[No article or Index item]
Black	1974	15	-	-	[No article or Index item]
Black	1984	15	-	-	[No article or Index item]
Black	1990	15	16	308	[No Article] (via: Christianity > The Church and its History > Church year > Liturgical colours)
Black	2005	15	16	308	[No Article] (via: Christianity > The Church and its History > Church year > Liturgical colours)
Black	2011	Web	-	EBchecked/topic/67435/black	black (colour)
Canada	1771	1	2	16	Canada
Canada	1778	2	3	1602-1604	Canada
Canada	1790	3	4	75-78	Canada
Canada	1798	3	-	-	[Data missing]
Canada	1842	7	6	55-69	Canada
Canada	1853	8	6	132-159	Canada
Canada	1878	9	4	765-782	Canada
Canada	1910	11	5	142-168	Canada
Canada	1968	14	4	728-760	Canada

Table 6 (continued)

Encyclopedia Britannica article corpus with bibliographic information

Article	Year	Edition	Volume	Pages	Actual Heading
Canada	1974	15	3	713-751	Canada
Canada	1984	15	3	713-751	Canada
Canada	1990	15	15	XXX	Canada
Canada	2005	15	15	440-513	Canada
Canada	2011	Web	-	EBchecked/topic/91513/Canada	Canada
Cannon	1771	1	2	24-26	Cannon
Cannon	1778	2	3	1625	Cannon
Cannon	1790	3	4	115-116	Cannon
Cannon	1798	3	4	115	Cannon
Cannon	1842	7	6	84-89	Cannon
Cannon	1853	8	6	177-186	Cannon
Cannon	1878	9	2	557	[No Article]
					(via: Arms and Armour > Artillery)
Cannon	1910	11	5	188-189	Cannon
Cannon	1968	14	4	791	Cannon
Cannon	1974	15	19	683	Advent of cannon
					(via: Weapons and Delivery Systems >
					Development of early weapons > Gunpowder
					weapons to Mid-19th Century > Artillery)
Cannon	1984	15	19	683	Advent of cannon (via: Weapons and
					Delivery Systems > Development of early
					weapons > Gunpowder weapons
					to Mid-19th Century > Artillery)
Cannon	1990	15	29	529-544	Enduring value of cast-bronze cannon
					(via: The Technology of War > Military
					technology before the modern era >
					The gunpowder revolution, c. 1300--1650 >
					The development of artillery)
Cannon	2005	15	29	555-559	Cannons (via: War, Technology of >
					Modern weapons and weapon systems
					> Artillery)
Cannon	2011	Web	-	EBchecked/topic/92758/cannon	cannon (weapon)
Cat	1771	1	2	45	Cat
Cat	1778	2	3	1720	Cat
Cat	1790	3	4	242-243	Cat
Cat	1798	3	4	242-243	Cat
Cat	1842	7	6	237	[No Article] (via: Cat-)
Cat	1853	8	7	168-171	Genus Felis, Linn. (via: Mammalia >
					Division III.- Carnivora. > Tribe II.-
					Digitigrada. > 3d Subdivision)
Cat	1878	9	5	205-206	Cat
Cat	1910	11	5	487-489	Cat
Cat	1968	14	5	52-54	Cat
Cat	1974	15	3	996-1000	Cat
Cat	1984	15	3	996-1000	Cat
Cat	1990	15	2	464-465	Cats (family Felidae)
					(via: Mammals > Major mammal orders >
					Carnivora > Survey of carnivore families)
Cat	2005	15	15	561-564	Cats, Domestic
Cat	2011	Web	-	EBchecked/topic/168536/domestic-cat	domestic cat (Mammal)
Democracy	1771	1	2	415	Democracy
Democracy	1778	2	4	2413	Democracy
Democracy	1790	3	5	749	Democracy
Democracy	1798	3	5	749	Democracy
Democracy	1842	7	7	708	Democracy
Democracy	1853	8	7	725	Democracy
Democracy	1878	9	9	10	[No article] (via: Government >
					Forms of Government > Three Standard
					Forms > Greek Politics)
Democracy	1910	11	8	1 - 2	Democracy
Democracy	1968	14	7	215-224	Democracy

Table 6 (continued)

Encyclopedia Britannica article corpus with bibliographic information

Article	Year	Edition	Volume	Pages	Actual Heading
Democracy	1974	15	14	715-716	Constitutional democracy (via: Political Systems > Forms of government > Types and models > The exercise and control of power)
Democracy	1984	15	14	715	Constitutional democracy (via: Political Systems > Forms of government > Types and models > The exercise and control of power)
Democracy	1990	15	4	991-993	Constitutional democracy (via: Political systems > Typologies of government > Governments classified by mode of succession > Autocratic versus nonautocratic rule)
Democracy	2005	15	17	226A-H	Democracy
Democracy	2011	Web	-	EBchecked/topic/157129/democracy	democracy
Human	1771	1	2	801	Human
Human	1778	2	5	3752	Human
Human	1790	3	8	708	Human
Human	1798	3	8	708	Human
Human	1842	7	11	720	Human
Human	1853	8	2	771-777	Human Anatomy (via: Anatomy)
Human	1878	9	1	819-908	Special anatomy of the Human Body. (via: Anatomy)
Human	1910	11	-	-	[No article or Index item]
Human	1968	14	11	827-828	Human Nature
Human	1974	15	8	1136-1168	Human Behaviour & Human Culture
Human	1984	15	8	1136-1168	Human Behaviour & Human Culture
Human	1990	15	14	708-722	The Development of Human Behaviour
Human	2005	15	18	838-843	Homo sapiens (via: Evolution, Human > Members of the family of Hominidae)
Human	2011	Web	-	EBchecked/topic/275376/human-being	human being
London	1771	1	2	1024	London
London	1778	2	6	4272	London
London	1790	3	10	232-272	London
London	1798	3	10	232-273	London
London	1842	7	13	485-549	London
London	1853	8	13	658-680	London
London	1878	9	14	818-852	London
London	1910	11	16	938-968	London
London	1968	14	14	263-290	London
London	1974	15	11	90-106	London
London	1984	15	11	90-106	London
London	1990	15	23	291-305	London
London	2005	15	23	283-298	London
London	2011	Web	-	EBchecked/topic/346821/London	London (England, United Kingdom)
Ocean	1771	1	3	409	Ocean
Ocean	1778	2	7	5459	Ocean
Ocean	1790	3	13	178	Ocean
Ocean	1798	3	13	178	Ocean
Ocean	1842	7	17	516-520	Ocean (via: Physical Geography > Distribution of the Waters of the Globe)
Ocean	1853	8	17	572-584	Of the sea (via: Physical Geography)
Ocean	1878	9	10	211	The Ocean or Water-Envelope of the Earth (via: Geography > Physical Geography)
Ocean	1910	11	19	967-987	Ocean/Oceanography
Ocean	1968	14	16	837-851	Ocean and Oceanography
Ocean	1974	15	13	482-504	Oceans and Seas
Ocean	1984	15	13	482-504	Oceans and Seas
Ocean	1990	15	25	123-195	Oceans
Ocean	2005	15	25	134-196	Oceans
Ocean	2011	Web	-	EBchecked/topic/424285/ocean	ocean (Earth Feature)

Table 7
Wikipedia article corpus

Article	EB	Ratio	Ed	Optimal Date	Actual title	Revision URL
America	1771	0.0000	A	2001-09-06	United_States	http://en.wikipedia.org/w/index.php?title=United_States&oldid=291561
America	1778	0.0300	B	2001-12-22	United_States	http://en.wikipedia.org/w/index.php?title=United_States&oldid=335284445
America	1790	0.0800	C	2002-06-17	United_States	http://en.wikipedia.org/w/index.php?title=United_States&oldid=99901
America	1798	0.1125	D	2002-10-11	United_States	http://en.wikipedia.org/w/index.php?title=United_States&oldid=354404
America	1842	0.2960	E	2004-07-25	United_States	http://en.wikipedia.org/w/index.php?title=United_States&oldid=4865179
America	1853	0.3420	F	2005-01-04	United_States	http://en.wikipedia.org/w/index.php?title=United_States&oldid=9111093
America	1878	0.4460	G	2006-01-09	United_States	http://en.wikipedia.org/w/index.php?title=United_States&oldid=34517027
America	1910	0.5790	H	2007-04-27	United_States	http://en.wikipedia.org/w/index.php?title=United_States&oldid=118382643
America	1968	0.8210	I	2009-09-03	United_States	http://en.wikipedia.org/w/index.php?title=United_States&oldid=311722899
America	1974	0.8460	J	2009-12-01	United_States	http://en.wikipedia.org/w/index.php?title=United_States&oldid=329166552
America	1984	0.8875	K	2010-04-28	United_States	http://en.wikipedia.org/w/index.php?title=United_States&oldid=352295565
America	1990	0.9125	L	2010-07-26	United_States	http://en.wikipedia.org/w/index.php?title=United_States&oldid=375599052
America	2005	0.9750	M	2011-03-05	United_States	http://en.wikipedia.org/w/index.php?title=United_States&oldid=417229472
America	2011	1.0000	N	2011-06-01	United_States	http://en.wikipedia.org/w/index.php?title=United_States&oldid=432073495
Atheism	1771	0.0000	A	2001-11-13	Atheism	http://en.wikipedia.org/w/index.php?title=Atheism&oldid=333968579
Atheism	1778	0.0300	B	2002-02-26	Atheism	http://en.wikipedia.org/w/index.php?title=Atheism&oldid=44182
Atheism	1790	0.0800	C	2002-08-19	Atheism	http://en.wikipedia.org/w/index.php?title=Atheism&oldid=354022
Atheism	1798	0.1125	D	2002-12-10	Atheism	http://en.wikipedia.org/w/index.php?title=Atheism&oldid=488422
Atheism	1842	0.2960	E	2004-09-10	Atheism	http://en.wikipedia.org/w/index.php?title=Atheism&oldid=5932251
Atheism	1853	0.3420	F	2005-02-18	Atheism	http://en.wikipedia.org/w/index.php?title=Atheism&oldid=10516157
Atheism	1878	0.4460	G	2006-02-16	Atheism	http://en.wikipedia.org/w/index.php?title=Atheism&oldid=39934744
Atheism	1910	0.5790	H	2007-05-26	Atheism	http://en.wikipedia.org/w/index.php?title=Atheism&oldid=133683781
Atheism	1968	0.8210	I	2009-09-16	Atheism	http://en.wikipedia.org/w/index.php?title=Atheism&oldid=315084288
Atheism	1974	0.8460	J	2009-12-12	Atheism	http://en.wikipedia.org/w/index.php?title=Atheism&oldid=331236393
Atheism	1984	0.8875	K	2010-05-06	Atheism	http://en.wikipedia.org/w/index.php?title=Atheism&oldid=360609650
Atheism	1990	0.9125	L	2010-08-01	Atheism	http://en.wikipedia.org/w/index.php?title=Atheism&oldid=376656644
Atheism	2005	0.9750	M	2011-03-07	Atheism	http://en.wikipedia.org/w/index.php?title=Atheism&oldid=417649563
Atheism	2011	1.0000	N	2011-06-01	Atheism	http://en.wikipedia.org/w/index.php?title=Atheism&oldid=431908365
Black	1771	0.0000	A	2001-08-13	Black	http://en.wikipedia.org/w/index.php?title=Black&oldid=238963
Black	1778	0.0300	B	2001-11-28	Black	http://en.wikipedia.org/w/index.php?title=Black&oldid=238964
Black	1790	0.0800	C	2002-05-26	Black	http://en.wikipedia.org/w/index.php?title=Black&oldid=88013
Black	1798	0.1125	D	2002-09-20	Black	http://en.wikipedia.org/w/index.php?title=Black&oldid=969025
Black	1842	0.2960	E	2004-07-08	Black	http://en.wikipedia.org/w/index.php?title=Black&oldid=4565924
Black	1853	0.3420	F	2004-12-19	Black	http://en.wikipedia.org/w/index.php?title=Black&oldid=8807047
Black	1878	0.4460	G	2005-12-27	Black	http://en.wikipedia.org/w/index.php?title=Black&oldid=32868482
Black	1910	0.5790	H	2007-04-17	Black	http://en.wikipedia.org/w/index.php?title=Black&oldid=123673389
Black	1968	0.8210	I	2009-08-30	Black	http://en.wikipedia.org/w/index.php?title=Black&oldid=310528740
Black	1974	0.8460	J	2009-11-28	Black	http://en.wikipedia.org/w/index.php?title=Black&oldid=328413101
Black	1984	0.8875	K	2010-04-25	Black	http://en.wikipedia.org/w/index.php?title=Black&oldid=357978559
Black	1990	0.9125	L	2010-07-24	Black	http://en.wikipedia.org/w/index.php?title=Black&oldid=375791230
Black	2005	0.9750	M	2011-03-05	Black	http://en.wikipedia.org/w/index.php?title=Black&oldid=417127040
Black	2011	1.0000	N	2011-06-01	Black	http://en.wikipedia.org/w/index.php?title=Black&oldid=432130498
Canada	1771	0.0000	A	2001-10-31	Canada	http://en.wikipedia.org/w/index.php?title=Canada&oldid=349145578
Canada	1778	0.0300	B	2002-02-13	Canada	http://en.wikipedia.org/w/index.php?title=Canada&oldid=38436
Canada	1790	0.0800	C	2002-08-07	Canada	http://en.wikipedia.org/w/index.php?title=Canada&oldid=144906
Canada	1798	0.1125	D	2002-11-29	Canada	http://en.wikipedia.org/w/index.php?title=Canada&oldid=463623
Canada	1842	0.2960	E	2004-09-01	Canada	http://en.wikipedia.org/w/index.php?title=Canada&oldid=5611696
Canada	1853	0.3420	F	2005-02-09	Canada	http://en.wikipedia.org/w/index.php?title=Canada&oldid=9982679
Canada	1878	0.4460	G	2006-02-08	Canada	http://en.wikipedia.org/w/index.php?title=Canada&oldid=38815983
Canada	1910	0.5790	H	2007-05-20	Canada	http://en.wikipedia.org/w/index.php?title=Canada&oldid=132310531
Canada	1968	0.8210	I	2009-09-13	Canada	http://en.wikipedia.org/w/index.php?title=Canada&oldid=313643728
Canada	1974	0.8460	J	2009-12-10	Canada	http://en.wikipedia.org/w/index.php?title=Canada&oldid=330902707
Canada	1984	0.8875	K	2010-05-04	Canada	http://en.wikipedia.org/w/index.php?title=Canada&oldid=360079157
Canada	1990	0.9125	L	2010-07-31	Canada	http://en.wikipedia.org/w/index.php?title=Canada&oldid=376225180
Canada	2005	0.9750	M	2011-03-06	Canada	http://en.wikipedia.org/w/index.php?title=Canada&oldid=417353526
Canada	2011	1.0000	N	2011-06-01	Canada	http://en.wikipedia.org/w/index.php?title=Canada&oldid=432027011
Cannon	1771	0.0000	A	2001-11-09	Cannon	http://en.wikipedia.org/w/index.php?title=Cannon&oldid=244392
Cannon	1778	0.0300	B	2002-02-22	Cannon	http://en.wikipedia.org/w/index.php?title=Cannon&oldid=11658
Cannon	1790	0.0800	C	2002-08-15	Cannon	http://en.wikipedia.org/w/index.php?title=Cannon&oldid=460974
Cannon	1798	0.1125	D	2002-12-07	Cannon	http://en.wikipedia.org/w/index.php?title=Cannon&oldid=832796
Cannon	1842	0.2960	E	2004-09-08	Cannon	http://en.wikipedia.org/w/index.php?title=Cannon&oldid=6022752

Table 7 (continued)

Wikipedia article corpus

Article	EB	Ratio	Ed	Optimal Date	Actual title	Revision URL
Cannon	1853	0.3420	F	2005-02-15	Cannon	http://en.wikipedia.org/w/index.php?title=Cannon&oldid=10704093
Cannon	1878	0.4460	G	2006-02-13	Cannon	http://en.wikipedia.org/w/index.php?title=Cannon&oldid=39935885
Cannon	1910	0.5790	H	2007-05-24	Cannon	http://en.wikipedia.org/w/index.php?title=Cannon&oldid=133183718
Cannon	1968	0.8210	I	2009-09-15	Cannon	http://en.wikipedia.org/w/index.php?title=Cannon&oldid=315591175
Cannon	1974	0.8460	J	2009-12-11	Cannon	http://en.wikipedia.org/w/index.php?title=Cannon&oldid=331129704
Cannon	1984	0.8875	K	2010-05-05	Cannon	http://en.wikipedia.org/w/index.php?title=Cannon&oldid=360227756
Cannon	1990	0.9125	L	2010-07-31	Cannon	http://en.wikipedia.org/w/index.php?title=Cannon&oldid=374629379
Cannon	2005	0.9750	M	2011-03-07	Cannon	http://en.wikipedia.org/w/index.php?title=Cannon&oldid=417862063
Cannon	2011	1.0000	N	2011-06-01	Cannon	http://en.wikipedia.org/w/index.php?title=Cannon&oldid=431915178
Cat	1771	0.0000	A	2001-11-09	Cat	http://en.wikipedia.org/w/index.php?title=Cat&oldid=331658056
Cat	1778	0.0300	B	2002-02-22	Cat	http://en.wikipedia.org/w/index.php?title=Cat&oldid=34498
Cat	1790	0.0800	C	2002-08-15	Cat	http://en.wikipedia.org/w/index.php?title=Cat&oldid=176303
Cat	1798	0.1125	D	2002-12-07	Cat	http://en.wikipedia.org/w/index.php?title=Cat&oldid=617951
Cat	1842	0.2960	E	2004-09-08	Cat	http://en.wikipedia.org/w/index.php?title=Cat&oldid=5791849
Cat	1853	0.3420	F	2005-02-15	Cat	http://en.wikipedia.org/w/index.php?title=Cat&oldid=10333032
Cat	1878	0.4460	G	2006-02-13	Cat	http://en.wikipedia.org/w/index.php?title=Cat&oldid=39505072
Cat	1910	0.5790	H	2007-05-24	Cat	http://en.wikipedia.org/w/index.php?title=Cat&oldid=133080530
Cat	1968	0.8210	I	2009-09-15	Cat	http://en.wikipedia.org/w/index.php?title=Cat&oldid=314119231
Cat	1974	0.8460	J	2009-12-11	Cat	http://en.wikipedia.org/w/index.php?title=Cat&oldid=331058169
Cat	1984	0.8875	K	2010-05-05	Cat	http://en.wikipedia.org/w/index.php?title=Cat&oldid=359899910
Cat	1990	0.9125	L	2010-07-31	Cat	http://en.wikipedia.org/w/index.php?title=Cat&oldid=376401226
Cat	2005	0.9750	M	2011-03-07	Cat	http://en.wikipedia.org/w/index.php?title=Cat&oldid=418193284
Cat	2011	1.0000	N	2011-06-01	Cat	http://en.wikipedia.org/w/index.php?title=Cat&oldid=432181054
Democracy	1771	0.0000	A	2001-10-14	Democracy	http://en.wikipedia.org/w/index.php?title=Democracy&oldid=336610238
Democracy	1778	0.0300	B	2002-01-28	Democracy	http://en.wikipedia.org/w/index.php?title=Democracy&oldid=246500
Democracy	1790	0.0800	C	2002-07-22	Democracy	http://en.wikipedia.org/w/index.php?title=Democracy&oldid=168871
Democracy	1798	0.1125	D	2002-11-14	Democracy	http://en.wikipedia.org/w/index.php?title=Democracy&oldid=433165
Democracy	1842	0.2960	E	2004-08-20	Democracy	http://en.wikipedia.org/w/index.php?title=Democracy&oldid=5435913
Democracy	1853	0.3420	F	2005-01-29	Democracy	http://en.wikipedia.org/w/index.php?title=Democracy&oldid=9862444
Democracy	1878	0.4460	G	2006-01-30	Democracy	http://en.wikipedia.org/w/index.php?title=Democracy&oldid=37280541
Democracy	1910	0.5790	H	2007-05-13	Democracy	http://en.wikipedia.org/w/index.php?title=Democracy&oldid=130912557
Democracy	1968	0.8210	I	2009-09-10	Democracy	http://en.wikipedia.org/w/index.php?title=Democracy&oldid=313048773
Democracy	1974	0.8460	J	2009-12-07	Democracy	http://en.wikipedia.org/w/index.php?title=Democracy&oldid=330185778
Democracy	1984	0.8875	K	2010-05-02	Democracy	http://en.wikipedia.org/w/index.php?title=Democracy&oldid=359694337
Democracy	1990	0.9125	L	2010-07-29	Democracy	http://en.wikipedia.org/w/index.php?title=Democracy&oldid=376065264
Democracy	2005	0.9750	M	2011-03-06	Democracy	http://en.wikipedia.org/w/index.php?title=Democracy&oldid=417473084
Democracy	2011	1.0000	N	2011-06-01	Democracy	http://en.wikipedia.org/w/index.php?title=Democracy&oldid=432047125
Human	1771	0.0000	A	2001-10-03	Human	http://en.wikipedia.org/w/index.php?title=Human&oldid=334736650
Human	1778	0.0300	B	2002-01-17	Human	http://en.wikipedia.org/w/index.php?title=Human&oldid=3745464
Human	1790	0.0800	C	2002-07-12	Human	http://en.wikipedia.org/w/index.php?title=Human&oldid=3745484
Human	1798	0.1125	D	2002-11-04	Human	http://en.wikipedia.org/w/index.php?title=Human&oldid=3745518
Human	1842	0.2960	E	2004-08-13	Human	http://en.wikipedia.org/w/index.php?title=Human&oldid=5206772
Human	1853	0.3420	F	2005-01-22	Human	http://en.wikipedia.org/w/index.php?title=Human&oldid=9606672
Human	1878	0.4460	G	2006-01-24	Human	http://en.wikipedia.org/w/index.php?title=Human&oldid=36562058
Human	1910	0.5790	H	2007-05-08	Human	http://en.wikipedia.org/w/index.php?title=Human&oldid=129186830
Human	1968	0.8210	I	2009-09-08	Human	http://en.wikipedia.org/w/index.php?title=Human&oldid=312511143
Human	1974	0.8460	J	2009-12-06	Human	http://en.wikipedia.org/w/index.php?title=Human&oldid=329757203
Human	1984	0.8875	K	2010-05-01	Human	http://en.wikipedia.org/w/index.php?title=Human&oldid=360595142
Human	1990	0.9125	L	2010-07-28	Human	http://en.wikipedia.org/w/index.php?title=Human&oldid=375955172
Human	2005	0.9750	M	2011-03-06	Human	http://en.wikipedia.org/w/index.php?title=Human&oldid=416937943
Human	2011	1.0000	N	2011-06-01	Human	http://en.wikipedia.org/w/index.php?title=Human&oldid=432033632
London	1771	0.0000	A	2001-10-08	London	http://en.wikipedia.org/w/index.php?title=London&oldid=331657753
London	1778	0.0300	B	2002-01-22	London	http://en.wikipedia.org/w/index.php?title=London&oldid=263884
London	1790	0.0800	C	2002-07-17	London	http://en.wikipedia.org/w/index.php?title=London&oldid=127948
London	1798	0.1125	D	2002-11-08	London	http://en.wikipedia.org/w/index.php?title=London&oldid=418301
London	1842	0.2960	E	2004-08-16	London	http://en.wikipedia.org/w/index.php?title=London&oldid=5386121
London	1853	0.3420	F	2005-01-25	London	http://en.wikipedia.org/w/index.php?title=London&oldid=9657137
London	1878	0.4460	G	2006-01-27	London	http://en.wikipedia.org/w/index.php?title=London&oldid=36980911
London	1910	0.5790	H	2007-05-10	London	http://en.wikipedia.org/w/index.php?title=London&oldid=130108545
London	1968	0.8210	I	2009-09-09	London	http://en.wikipedia.org/w/index.php?title=London&oldid=312633659
London	1974	0.8460	J	2009-12-06	London	http://en.wikipedia.org/w/index.php?title=London&oldid=329957634

Table 7 (continued)

Wikipedia article corpus

Article	EB	Ratio	Ed	Optimal Date	Actual title	Revision URL
London	1984	0.8875	K	2010-05-02	London	http://en.wikipedia.org/w/index.php?title=London&oldid=359703505
London	1990	0.9125	L	2010-07-29	London	http://en.wikipedia.org/w/index.php?title=London&oldid=376094221
London	2005	0.9750	M	2011-03-06	London	http://en.wikipedia.org/w/index.php?title=London&oldid=417305498
London	2011	1.0000	N	2011-06-01	London	http://en.wikipedia.org/w/index.php?title=London&oldid=431503481
Ocean	1771	0.0000	A	2001-06-18	Ocean	http://en.wikipedia.org/w/index.php?title=Ocean&oldid=271538
Ocean	1778	0.0300	B	2001-10-05	Ocean	http://en.wikipedia.org/w/index.php?title=Ocean&oldid=331657006
Ocean	1790	0.0800	C	2002-04-05	Ocean	http://en.wikipedia.org/w/index.php?title=Ocean&oldid=39763
Ocean	1798	0.1125	D	2002-08-01	Ocean	http://en.wikipedia.org/w/index.php?title=Ocean&oldid=196287
Ocean	1842	0.2960	E	2004-05-29	Ocean	http://en.wikipedia.org/w/index.php?title=Ocean&oldid=4054829
Ocean	1853	0.3420	F	2004-11-13	Ocean	http://en.wikipedia.org/w/index.php?title=Ocean&oldid=7706679
Ocean	1878	0.4460	G	2005-11-26	Ocean	http://en.wikipedia.org/w/index.php?title=Ocean&oldid=29418009
Ocean	1910	0.5790	H	2007-03-24	Ocean	http://en.wikipedia.org/w/index.php?title=Ocean&oldid=117625833
Ocean	1968	0.8210	I	2009-08-20	Ocean	http://en.wikipedia.org/w/index.php?title=Ocean&oldid=309144102
Ocean	1974	0.8460	J	2009-11-19	Ocean	http://en.wikipedia.org/w/index.php?title=Ocean&oldid=326701256
Ocean	1984	0.8875	K	2010-04-19	Ocean	http://en.wikipedia.org/w/index.php?title=Ocean&oldid=350698493
Ocean	1990	0.9125	L	2010-07-19	Ocean	http://en.wikipedia.org/w/index.php?title=Ocean&oldid=372621644
Ocean	2005	0.9750	M	2011-03-03	Ocean	http://en.wikipedia.org/w/index.php?title=Ocean&oldid=416944978
Ocean	2011	1.0000	N	2011-06-01	Ocean	http://en.wikipedia.org/w/index.php?title=Ocean&oldid=431889533

Table 8

Bringhurst's sociological hierarchy of typefaces,

Bringhurst, 2002, p.54

Primary	roman lower case
Secondary	Roman Upper Case ROMAN SMALL CAPS roman text figures: 123 <i>italic lower case</i>
Tertiary	<i>True Italic (Cursive) Upper Case</i> <i>italic text figures: 123</i> SLOPED SMALL CAPS Roman Titling Figures: 123 bold lower case
Quaternary	<i>False Italic (Sloped Roman) Upper Case</i> Bold Upper Case BOLD SMALL CAPS bold text figures: 123 <i>bold italic lower case</i>
Quintary	<i>Italic Titling Figures: 123</i> <i>Bold Italic (Sloped Roman) Upper Case</i> <i>bold italic text figures: 123</i> Bold Titling figures: 123
Sextary	<i>Bold Italic Titling Figures: 123</i>

Table 9
Coding stylesheets

	Order	Font	Style	Case	Size	Position
EB-1771	Body	Caslon Serif	Roman	Sentence		Full-measure
H1	Heading level 1		ROMAN italic SC	ALLCAPS lowercase		Run-in
H2	Heading level 2		Smallcaps italic	Sentence		Run-in
EB-1778	Body	Caslon Serif	Roman	Sentence		Full-measure
H1	Heading level 1		ROMAN italic SC	ALLCAPS lowercase		Run-in
H2	Heading level 2		Smallcaps italic	Sentence		Run-in
H3	Heading level 3		# ¶ Roman	Sentence	-2pt	Shoulder
H4	Heading level 4		#. Italic	Noun		Run-in
H5	Heading level 5		#. Roman	Sentence		Run-in
EB-1790	Body	Caslon Serif	Roman	Sentence		Full-measure
H1	Heading level 1		ROMAN italic SC	ALLCAPS lowercase		Run-in
H2	Heading level 2		Smallcaps italic	Sentence		Run-in
H3	Heading level 3		## ¶ Roman	Sentence	-2pt	Shoulder
H4	Heading level 4		##. Italic Smallcaps	Roman. Noun		Run-in
H5	Heading level 5		##. Italic	Arabic. Sentence		Run-in
EB-1798	Body	Caslon Serif	Roman	Sentence		Full-measure
H1	Heading level 1		ROMAN italic SC	ALLCAPS lowercase		Run-in
H2	Heading level 2		Smallcapitalic	Sentence		Run-in
H3	Heading level 3		## ¶ Roman	Sentence	-2pt	Shoulder
H4	Heading level 4		##. Italic Smallcaps	Roman. Noun		Run-in
H5	Heading level 5		##. Italic	Arabic. Sentence		Run-in
H6	Heading level 6		##. Roman	Roman. Noun		Run-in
EB-1842	Body	Caslon Serif	Roman	Sentence case.		Full-measure
H1	Heading level 1		ROMAN	ALLCAPS		Run-in Crosshead
H2	Heading level 2		Smallcaps	Noun		Run-in
H3	Heading level 3		Smallcapsitalic	Noun		Run-in
H4	Heading level 4		Italic	Noun		Crosshead
GL	Glossa		Roman.	Sentence	-2pt	Shoulder
H5	Heading level 5		Smallcaps.–	Noun		Run-in
H6	Heading level 6		Italic.–	Sentence		Run-in
H7	Heading level 7		The Italic [...]	Noun		Run-in
EB-1853	Body	Caslon Serif	Roman	Sentence case.		Full-measure
H1	Heading level 1		ROMAN	ALLCAPS		Run-in Crosshead
H2	Heading level 2		Smallcaps	Noun		Run-in Crosshead
H3	Heading level 3		SC italic Italic	Noun		Run-in Crosshead
H4.1	Heading level 4.1		I. Roman	Noun		Shoulder
H4.2	Heading level 4.2		Sec. 1. Roman	Noun		Shoulder
H4	Heading level 4		Roman.	Sentence	-2pt	Shoulder
H5	Heading level 5		Italic.–	Sentence		Run-in
H6	Heading level 6		#[Roman]. Roman	Sentence		Run-in
EB-1878	Body	Caslon Serif	Roman	Sentence case.		Full-measure
H1	Heading level 1		ROMAN, letterspaced	ALLCAPS		Run-in Crosshead
H2	Heading level 2		ROMAN	ALLCAPS		Crosshead
H3	Heading level 3		SC	Noun		Run-in Crosshead
H4	Heading level 4		SCitalic Italic	Noun		Run-in Crosshead
H5	Heading level 5		Smallcaps.–	Sentence		Run-in
H6	Heading level 6		Italic.– SC.–	Sentence		Run-in
H7	Heading level 7		#. Roman a. Italic	Sentence		Run-in
H8	Heading level 8		Italic.–	Noun		Run-in
GL	Glossa		Roman. I. A.	Noun		Shoulder

Table 9 (continued)

Coding stylesheets

	Order	Font	Style	Case	Size	Position
EB-1910	Body	Scotch Serif	Roman	Sentence case.		Full-measure
H1	Heading level 1		BOLD,	ALLCAPS		Run-in
H2	Heading level 2		I. Smallcaps	Noun Caps.		Crosshead
H3	Heading level 3		I.-Italic	Noun Caps.		Crosshead
H4	Heading level 4		Smallcaps	Noun Caps.		Crosshead
H5	Heading level 5		Italic.-	Noun Caps.		Run-in
H6	Heading level 6		\$ #. Roman	Sentence case.		Run-in
H7	Heading level 7		#. / (a) Italic	Noun Caps.		Run-in
H8	Heading level 8		Bolditalic.	Sentence case.	-2pt	Shoulder / Inset
H9	Heading level 9		a. Roman / Italic.-	Sentence case.		Run-in
EB-1968	Body	Scotch	Roman	Sentence case.		Full-measure
H1	Heading level 1	Garamond	BOLD	ALLCAPS	+2pt	Run-in
H2	Heading level 2 diff		I. BOLD	ALLCAPS		Crosshead
H3	Heading level 3		A. Smallcaps	Noun Caps.		Crosshead
H4	Heading level 4 (US)		Italic	Noun Caps.		Crosshead
H5	Heading level 5 (US)		Smallcaps	Noun Caps.		Crosshead
H6	Heading level 6 diff		1. Bold.-	Noun Caps.		Run-in
H7	Heading level 7		Italic.-	Noun Caps.		Run-in
H7.1	Heading level 7.1		Roman :	Noun Caps.		Run-in
EB-1974	Body	Times New Roman	Roman	Sentence case.		Full-measure
H1	Heading level 1		Bold	Noun Case.	14pt	Sidehead
H2	Heading level 2		I. Bold	Sentence case.		Sidehead
H3	Heading level 3		UNDERLINED	ALLCAPS	-2pt	Sidehead
H4	Heading level 4		Bold.	Sentence case.		Run-in
H5	Heading level 5		Italic.	Sentence case.		Run-in
H6	Heading level 6		Roman:	Sentence case.		Run-in
GL	Glossa		Roman	Sentence case.		margin
EB-1984	Body	Nimbus Roman No 9	Roman	Sentence case.	11pt	Full-measure
H1	Heading level 1		Roman	Noun Case.	24pt	Crosshead
H2	Heading level 2		Bold	Allcaps	18pt	Crosshead
H3	Heading level 3		Bold	sentence case	14pt	Side-head
H4	Heading level 4		underline	Allcaps	-2pts	Side-head
H5	Heading level 5		Bold. [2emspace]	Sentence case.		run-in
H6	Heading level 6		Italic [2emspace]	Sentence case.		run-in
H7	Heading Level 7		(#) Roman	Sentence case.		run-in
H8	Heading level 8		Bold. [2emspace]	Sentence case.		run-in
H9	Heading level 9		Italic: [2emspaces]	Sentence case.		run-in
H10	Heading level 10		(Italic)	Sentence case.		run-in
GL	Glossa		Semibold	Sentence case.		margin

Table 9 (continued)

Coding stylesheets

	Order	Font	Style	Case	Size	Position
EB-1990	Body	Nimbus Roman No 9	Roman	Sentence case.	11pt	Full-measure
H1	Heading level 1		Roman	Noun Case.	24pt	Crosshead
H2	Heading level 2		Roman	Allcaps	18pt	Crosshead
H3	Heading level 3		Bold	sentence case	14pt	Side-head
H4	Heading level 4		underline	Allcaps	-2pt	Side-head
H5	Heading level 5		Bold. [2emspace]	Sentence case.		run-in
H6	Heading level 6		Italic [2emspace]	Sentence case.		run-in
H7	Heading Level 7		(#) Roman	Sentence case.		run-in
H8	Heading level 8		Bold. [2emspace]	Sentence case.		run-in
H9	Heading level 9		Italic: [2emspaces]	Sentence case.		run-in
H10	Heading level 10		Italic)	Sentence case.		run-in
GL	Glossa		Semibold	Sentence case.		margin
EB-2005	Body	Nimbus Roman No 9	Roman	Sentence case.	11pt	Full-measure
H1	Heading level 1		Roman	Noun Case.	24pt	Crosshead
H2	Heading level 2		Bold	Allcaps	18pt	Crosshead
H3	Heading level 3		Bold	sentence case	14pt	Side-head
H4	Heading level 4		underline	Allcaps	-2pt	Side-head
H5	Heading level 5		Bold. [2emspace]	Sentence case.		run-in
H6	Heading level 6		Italic [2emspace]	Sentence case.		run-in
H7	Heading Level 7		(#) Roman	Sentence case.		run-in
H8	Heading level 8		Bold. [2emspace]	Sentence case.		run-in
H9	Heading level 9		Italic: [2emspaces]	Sentence case.		run-in
H10	Heading level 10		(Italic)	Sentence case.		run-in
GL	Glossa		Semibold	Sentence case.		margin
EB-2011	Body	Georgia	Roman	Sentence case.	0.85em	Full-measure
H1	Heading level 1	Arial	Roman	Noun Case.	28px	Sidehead
H2.1	Heading level 1		Bold	Allcaps	11px	Sidehead
H2	Heading level 2		Roman 100% rule	Sentence case.	1.4em	Sidehead
H3	Heading level 3		Bold	Sentence case.	1.2em	Sidehead
H4	Heading level 4		Roman underline	ALLCAPS	1em	Sidehead
H5	Heading level 5		Bold	Noun Case.	1em	Sidehead
H6	Heading level 6		Italic	Sentence case.	1em	Sidehead
Wikipedia	Body	Georgia	Roman	Sentence	11.20 px	Full-measure
H1	Heading level 1	Sans-Serif	Roman	Noun	22.40 px	Sidehead
H2	Heading level 2		Roman	Noun	21.00 px	Sidehead
H3	Heading level 3		Bold	Noun	17.25 px	Sidehead
H4	Heading level 4		Bold	Noun	116%;	Sidehead
H5	Heading level 5		Bold	Noun	12.32 px	Sidehead
H6 (h5)	Heading level 6		Roman	Noun	12.32 px	Sidehead
H7	Heading level 7	Georgia	Italic	Noun		Sidehead
H8 (H6)	Heading level 7	Georgia	Italic	Sentence		Run-in

Figure 4
Sum of total heads per article
Encyclopedia Britannica

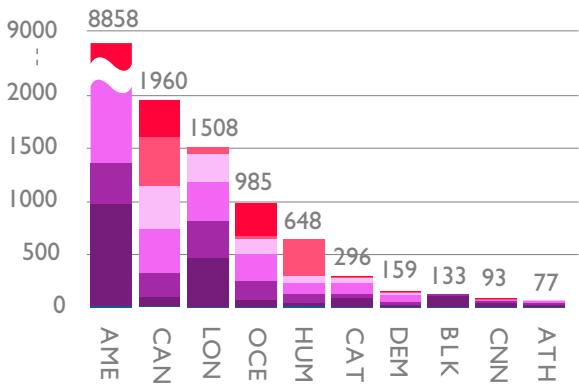


Figure 5
Sum of total heads per article
Wikipedia

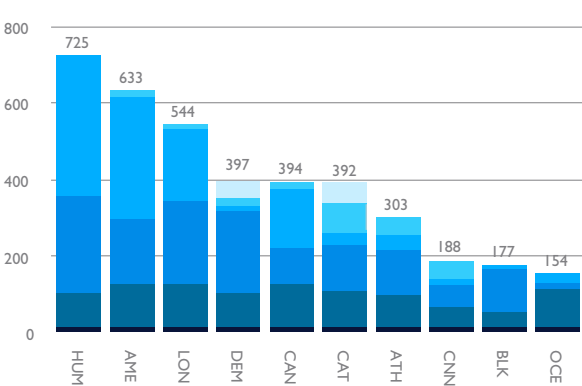


Figure 6
Sum of total heads per edition
Encyclopedia Britannica

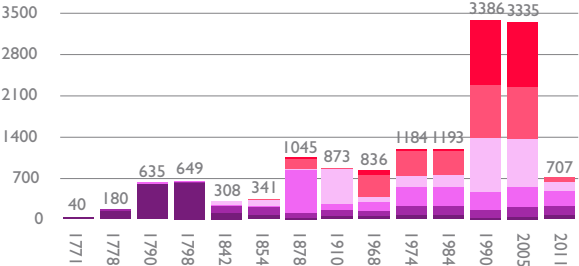


Figure 7
Sum of total heads per edition
Wikipedia

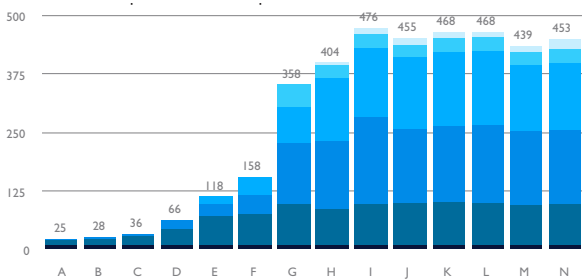


Figure 8

Chronology of heads per article for Britannica, segmented by rank



Figure 9

Chronology of heads per article for Wikipedia, segmented by rank

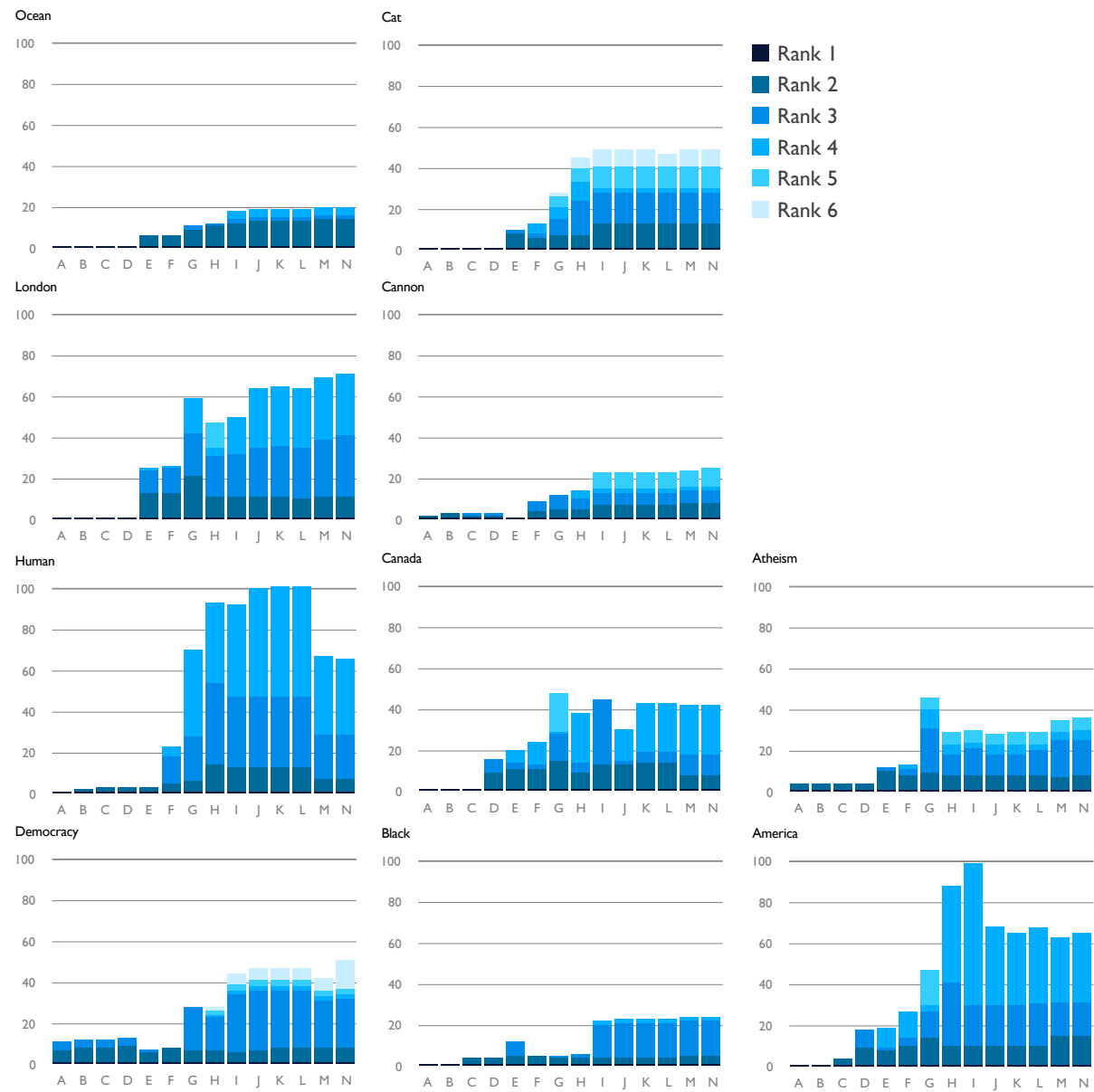
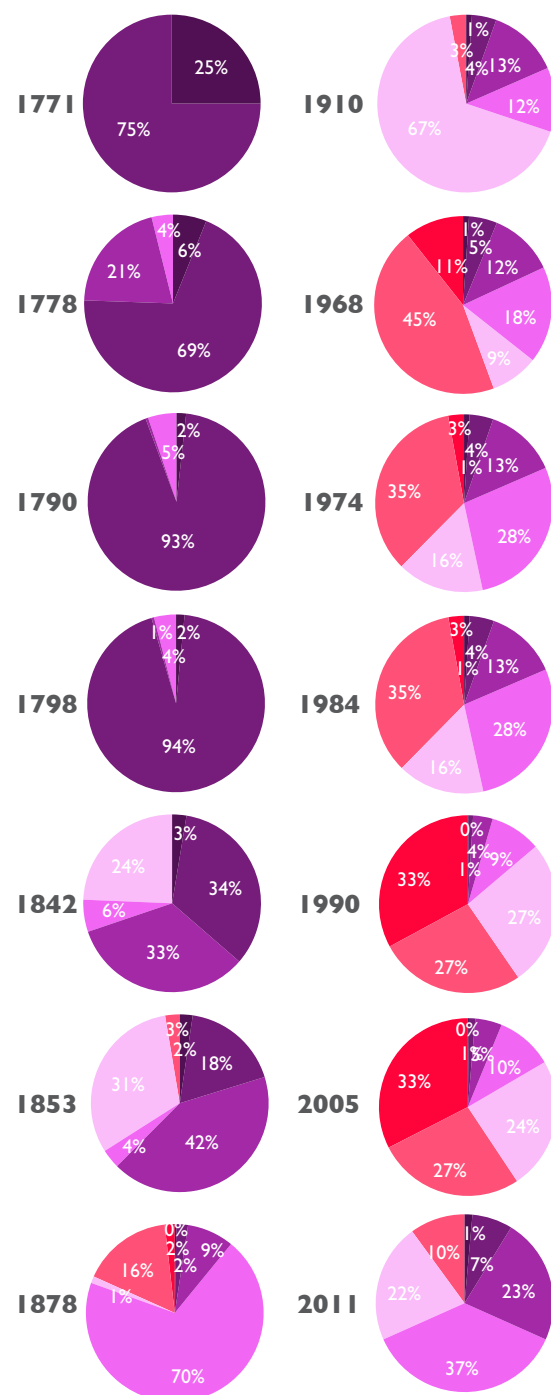


Figure 10

Percentage of ranks used within each edition

Encyclopedia Britannica

**Figure 11**

Percentage of ranks used within each edition

Wikipedia

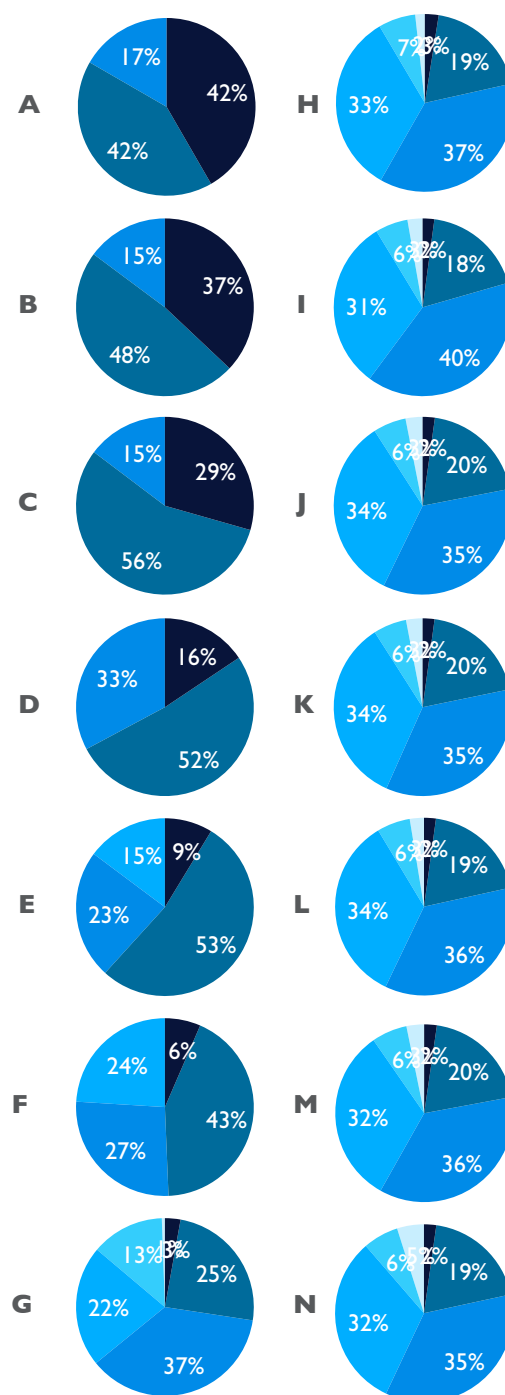


Table 10

Ranking by frequency of top 120 common words (orded by EB)

Word	EB	WK	Word	EB	WK
war	342	66	local	44	20
life	312	22	relations	43	26
government	292	90	exploration	42	18
settlement	264	2	against	42	1
social	225	9	federal	41	12
cultural	215	5	colonies	40	2
education	196	37	20th	40	20
patterns	194	7	toward	38	1
history	182	180	impact	38	11
economy	177	73	waves	37	1
physical	172	14	public	37	2
climate	170	41	higher	37	1
agriculture	147	2	constitutional	37	13
transportation	151	16	areas	37	2
geography	143	58	quebec	36	1
health	129	29	rights	35	4
general	117	4	origin	35	3
development	112	11	european	35	19
early	110	48	parties	34	9
great	109	7	temperature	33	2
economic	103	4	provincial	33	2
state	100	19	east	33	5
population	96	19	problems	32	1
world	95	35	peace	32	1
national	95	11	modern	32	38
system	94	2	military	32	34
political	93	31	cities	32	17
french	90	1	society	31	39
city	84	2	properties	31	7
water	83	6	executive	31	2
south	83	1	under	30	4
british	82	5	law	30	24
civil	80	13	expansion	30	12
west	77	8	confederation	30	12
other	75	13	common	30	1
century	72	15	church	30	1
foreign	68	45	characteristics	30	8
culture	66	99	tourism	28	6
party	65	1	types	27	16
arts	64	16	railways	27	5
politics	63	73	literature	27	39
policy	63	10	industrial	27	2
revolution	61	14	groups	27	13
north	60	1	rise	26	2
royal	58	8	religion	26	77
trade	57	17	movement	26	7
effects	57	12	independence	26	6
ethnic	56	26	anti	26	1
england	56	1	administrative	26	1
distribution	56	1	white	25	18
regions	55	7	governments	25	4
reconstruction	55	3	spanish	24	1
period	53	18	energy	24	14
systems	52	3	theatre	23	7
major	52	1	range	23	7
domestic	50	3	middle	23	28
immigration	49	4	evolution	23	36
role	47	8	depression	23	7
era	46	7	international	22	7
religious	45	5	history	182	180
reform	45	1	culture	66	99

Table 11

Ranking by frequency of top 120 common words (orded by WK)

Word	WK	EB	Word	WK	EB
government	292	90	transportation	151	16
religion	26	77	arts	64	16
economy	177	73	types	27	16
politics	63	73	income	14	16
war	342	66	popular	10	16
geography	143	58	languages	3	16
music	18	56	demography	2	16
list	2	56	sport	1	16
early	110	48	century	72	15
foreign	68	45	environment	16	15
technology	4	43	senses	8	15
climate	170	41	artillery	2	15
society	31	39	physical	172	14
literature	27	39	revolution	61	14
genetics	5	39	energy	24	14
modern	32	38	forces	22	14
language	10	38	19th	21	14
education	196	37	divisions	19	14
evolution	23	36	cold	17	14
world	95	35	film	8	14
science	4	35	elections	5	14
military	32	34	18th	4	14
arguments	10	34	operation	2	14
political	93	31	habitat	2	14
anatomy	9	31	direct	2	14
health	129	29	classical	2	14
middle	23	28	ethnicity	1	14
sports	17	27	civil	80	13
ethnic	56	26	other	75	13
relations	43	26	constitutional	37	13
ages	2	26	groups	27	13
law	30	24	museums	16	13
art	21	24	districts	8	13
motivation	2	24	television	6	13
emotion	1	24	representative	6	13
consciousness	1	24	ecology	2	13
love	2	23	domesticated	1	13
usage	1	23	cycle	1	13
life	312	22	effects	57	12
parks	17	22	federal	41	12
ancient	9	22	expansion	30	12
sleep	1	22	confederation	30	12
territories	15	21	marine	15	12
provinces	11	21	republic	14	12
race	8	21	opposition	14	12
centuries	3	21	geology	14	12
local	44	20	contemporary	11	12
20th	40	20	media	8	12
state	100	19	exclusion	6	12
population	96	19	crime	4	12
european	35	19	absorption	3	12
self	7	19	antiquity	2	12
period	53	18	thought	1	12
exploration	42	18	development	112	11
white	25	18	national	95	11
transport	16	18	impact	38	11
light	7	18	transition	15	11
expressions	1	18	rule	15	11
trade	57	17	peoples	14	11
cities	32	17	origins	14	11
economics	2	17	class	12	11

Figure 12

Comparison of the most frequent words in the Encyclopedia Britannica and Wikipedia

Highlighting the words that are the most frequent in both

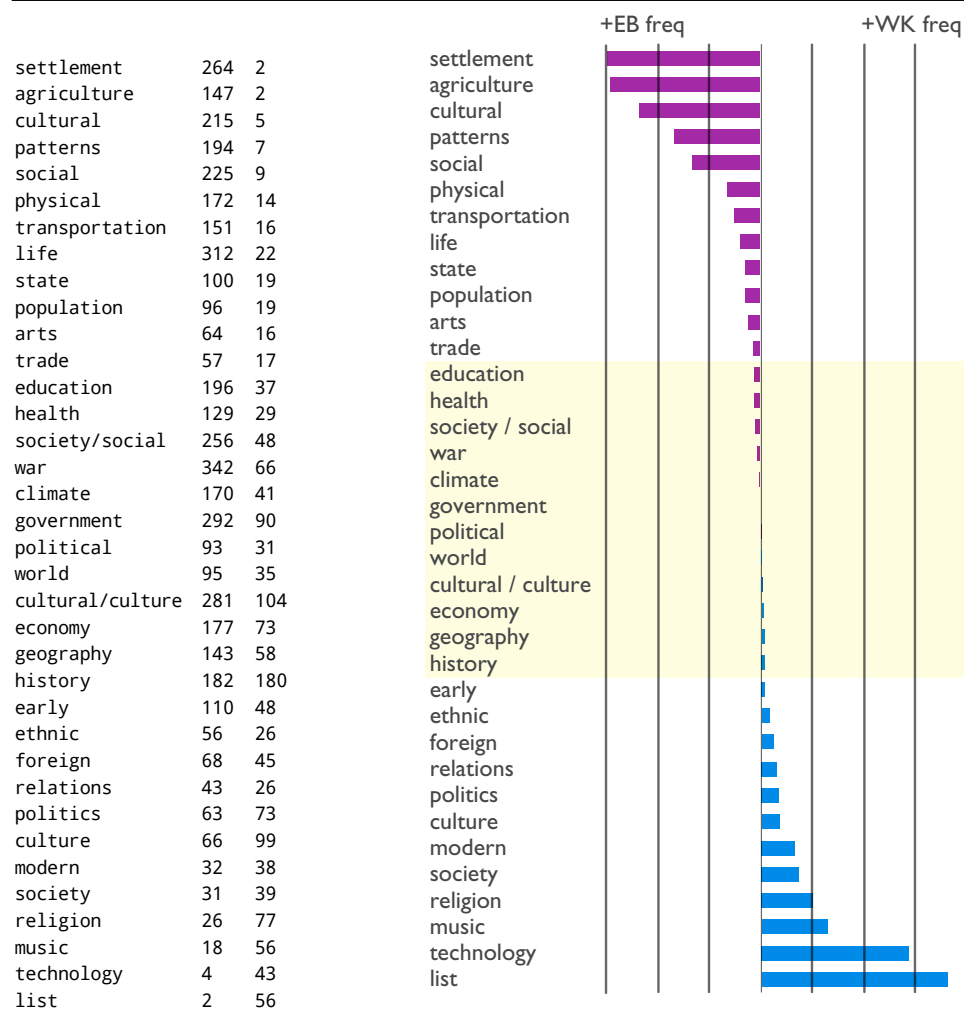


Table 12

Unique words

Encyclopedia Britannica

administration	202	different	32
land	196	seawater	32
conditions	161	silver	32
growth	144	tariff	32
industry	138	article	31
relief	135	lakes	31
animal	132	market	31
welfare	132	mountains	31
plant	112	natural	31
resources	104	republican	31
sea	100	army	30
act	97	colony	30
drainage	96	influence	30
ward	78	islands	30
soils	77	labour	30
indian	76	plan	30
first	75	second	30
mining	71	forest	29
power	62	street	29
statehood	60	surface	29
congress	58	town	29
indians	54	colleges	28
urban	53	features	28
between	51	finance	28
affairs	50	postwar	28
currents	49	universities	28
pacific	49	westminster	28
rural	49	1850	27
mountain	47	division	27
colonial	46	gulf	27
commerce	46	internal	27
fishing	46	manufacturing	27
southern	46	mineral	27
forestry	45	organization	27
character	44	powers	27
region	43	1763	26
defeated	41	britain	26
house	40	company	26
services	40	constitution	26
institutions	39	rebellion	26
western	39	acts	25
boston	38	bill	25
changes	38	clevelands	25
formation	38	convention	25
trends	38	cultures	25
court	37	fire	25
production	37	georgia	25
treaty	37	ice	25
york	36	nation	25
continental	35	president	25
lord	35	washington	25
virginia	35	1920	24
island	34	atlantic	24
coast	33	cannon	24
deal	33	climatic	24
gold	33	colour	24
minerals	33	democratic	24
northern	33	mississippi	24
plains	33	progressive	24
slavery	33	etymology	48
territory	33	biology	37

Table 13

Unique words

Wikipedia

philosophy	33	socialist	7
physiology	30	spoken	7
sexuality	29	tool	7
civilization	23	tuxedo	7
psychology	22	accent	6
gender	18	ambiguity	6
spirituality	18	beliefs	6
varieties	17	breed	6
vs	16	buses	6
explicit	15	clothing	6
feral	15	deception	6
implicit	15	deceptive	6
symbolism	14	electromagnetic	6
cobby	13	fiction	6
diet	13	glow	6
play	13	holidays	6
positive	13	ideologies	6
autocannon	12	ideology	6
colloquial	12	indoor	6
criticism	12	islamic	6
dualism	12	muslim	6
paleolithic	12	negative	6
semi	12	ontological	6
vote	12	policing	6
cradle	11	radical	6
featuring	11	republicanism	6
neolithic	11	secrecy	6
shot	11	seriousness	6
extraterrestrial	10	sociability	6
belief	9	supranational	6
bicolor	9	symbolic	6
1992	8	taxonomy	6
behavior	8	toys	6
beings	8	twin	6
cinema	8	twinned	6
consensus	8	annual	5
cuisine	8	apatheism	5
films	8	depictions	5
pigments	8	facts	5
sortition	8	rationale	5
terminology	8	tallest	5
aggregative	7	today	5
anarchist	7	topics	5
chambers	7	toponymy	5
colorpoint	7	trams	5
cosmopolitan	7	alternative	4
deliberative	7	alternatives	4
egypt	7	buying	4
epistemological	7	cycling	4
fashion	7	decay	4
feeding	7	facade	4
hinduism	7	reductionary	4
indic	7	slow	4
logical	7	termism	4
material	7	theater	4
monarchs	7	timeline	4
participatory	7	unsustainability	4
polity	7	volatility	4
prey	7	anthropocentric	3
rankings	7	belief	3
scratching	7	coordinates	3

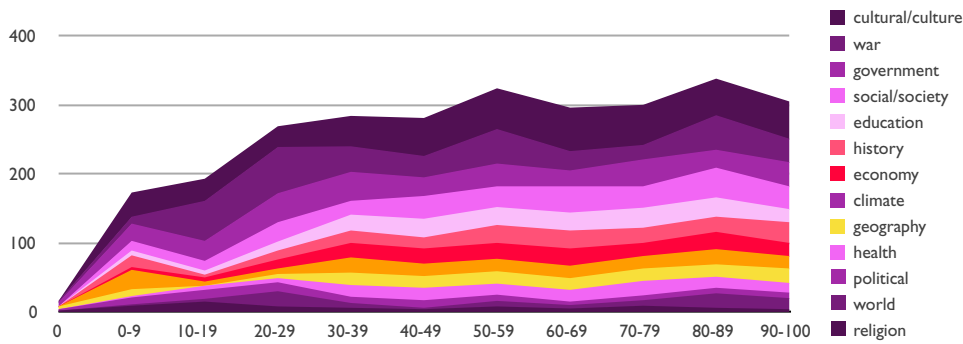
Table 14

RANK: Rank frequency per encyclopedia per word type (common/unique)

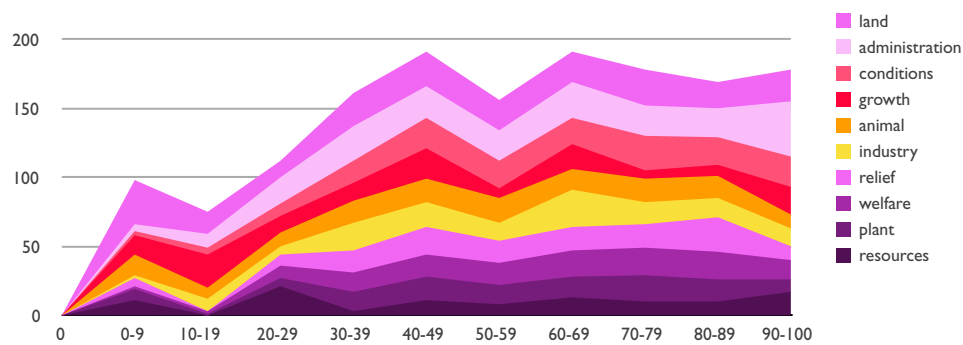
	0	1	2	3	4	5	6	7	SUM	0.5	0.33	0.25	0.2	0.166	0.143	SCORE
Britannica Common																
government	6	0	23	21	32	41	143	61	327	11.50	6.93	8.00	8.20	24.31	8.72	67.66
political	2	0	3	11	15	26	17	21	95	1.50	3.63	3.75	5.20	2.89	3.00	19.97
climate	0	0	15	10	16	5	117	13	176	7.50	3.30	4.00	1.00	19.89	1.86	37.55
world	0	0	8	16	28	10	17	16	95	4.00	5.28	7.00	2.00	2.89	2.29	23.46
economy	0	0	11	14	7	130	2	13	177	5.50	4.62	1.75	26.00	0.34	1.86	40.07
war	2	0	24	34	86	122	62	58	388	12.00	11.22	21.50	24.40	10.54	8.29	87.95
geography	4	0	11	2	128	1	1	0	147	5.50	0.66	32.00	0.20	0.17	0.00	38.53
history	0	4	34	2	133	16	1	4	194	17.00	0.66	33.25	3.20	0.17	0.57	54.85
health	0	0	5	6	8	4	106	4	133	2.50	1.98	2.00	0.80	18.02	0.57	25.87
education	0	0	3	13	19	9	122	38	204	1.50	4.29	4.75	1.80	20.74	5.43	38.51
cultural/culture	0	4	18	34	34	156	156	81	483	9.00	11.22	8.50	31.20	26.52	11.58	98.02
social/society	0	0	12	27	22	159	47	17	284	6.00	8.91	5.50	31.80	7.99	2.43	62.63
religion	2	0	5	10	20	11	15	13	76	2.50	3.30	5.00	2.20	2.55	1.86	17.41
Wikipedia Common																
government	0	0	20	43	40	1	0	0	104	10.00	14.19	10.00	0.20	0.00	0.00	34.39
political	0	0	9	12	9	1	0	0	31	4.50	3.96	2.25	0.20	0.00	0.00	10.91
climate	0	0	19	8	13	1	0	0	41	9.50	2.64	3.25	0.20	0.00	0.00	15.59
world	0	0	5	15	9	0	0	0	29	2.50	4.95	2.25	0.00	0.00	0.00	9.70
economy	0	0	40	9	22	2	0	0	73	20.00	2.97	5.50	0.40	0.00	0.00	28.87
war	0	0	4	33	29	0	0	0	66	2.00	10.89	7.25	0.00	0.00	0.00	20.14
geography	0	0	29	3	23	3	0	0	58	14.50	0.99	5.75	0.60	0.00	0.00	21.84
history	0	0	88	25	46	18	10	0	187	44.00	8.25	11.50	3.60	1.70	0.00	69.05
health	0	0	8	11	4	0	6	0	29	4.00	3.63	1.00	0.00	1.02	0.00	9.65
education	0	0	11	9	17	2	0	0	39	5.50	2.97	4.25	0.40	0.00	0.00	13.12
cultural/culture	0	0	50	14	34	3	5	0	106	25.00	4.62	8.50	0.60	0.85	0.00	39.57
social/society	0	0	13	12	23	0	0	0	48	6.50	3.96	5.75	0.00	0.00	0.00	16.21
religion	0	0	16	46	27	4	0	0	93	8.00	15.18	6.75	0.80	0.00	0.00	30.73
Britannica Unique																
administration	0	0	11	12	26	145	15	4	213	5.50	3.96	6.50	29.00	2.55	0.57	48.08
land	0	0	13	8	15	146	26	13	221	6.50	2.64	3.75	29.20	4.42	1.86	48.37
conditions	0	0	10	9	8	130	1	3	161	5.00	2.97	2.00	26.00	0.17	0.43	36.57
growth	0	0	4	10	29	54	16	31	144	2.00	3.30	7.25	10.80	2.72	4.43	30.50
industry	0	0	2	3	7	7	104	15	138	1.00	0.99	1.75	1.40	17.68	2.15	24.97
relief	0	0	5	3	4	5	115	3	135	2.50	0.99	1.00	1.00	19.55	0.43	25.47
animal	0	0	14	5	10	2	109	2	142	7.00	1.65	2.50	0.40	18.53	0.29	30.37
welfare	0	0	0	5	7	5	112	3	132	0.00	1.65	1.75	1.00	19.04	0.43	23.87
plant	0	0	1	5	4	3	103	3	119	0.50	1.65	1.00	0.60	17.51	0.43	21.69
resources	0	0	0	13	18	18	53	2	104	0.00	4.29	4.50	3.60	9.01	0.29	21.69
Wikipedia Unique																
etymology	0	0	45	2	1	0	0	0	48	22.50	0.66	0.25	0.00	0.00	0.00	23.41
biology	0	0	20	0	17	0	0	0	37	10.00	0.00	4.25	0.00	0.00	0.00	14.25
philosophy	0	0	1	21	11	0	0	0	33	0.50	6.93	2.75	0.00	0.00	0.00	10.18
physiology	0	0	7	13	10	0	0	0	30	3.50	4.29	2.50	0.00	0.00	0.00	10.29
sexuality	0	0	0	18	11	0	0	0	29	0.00	5.94	2.75	0.00	0.00	0.00	8.69
civilization	0	0	0	12	11	0	0	0	23	0.00	3.96	2.75	0.00	0.00	0.00	6.71
gender	0	0	0	11	7	0	0	0	18	0.00	3.63	1.75	0.00	0.00	0.00	5.38
psychology	0	0	12	1	9	0	0	0	22	6.00	0.33	2.25	0.00	0.00	0.00	8.58
spirituality	0	0	1	9	8	0	0	0	18	0.50	2.97	2.00	0.00	0.00	0.00	5.47
varieties	0	0	3	7	0	0	7	0	17	1.50	2.31	0.00	0.00	1.19	0.00	5.00

Figure 13
SEQUENCE: Frequency of article position

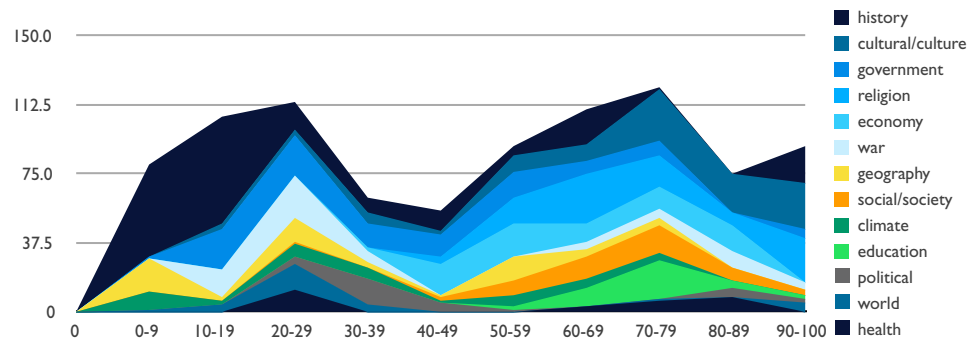
Encyclopedia Britannica (common)



Encyclopedia Britannica (unique)



Wikipedia (common)



Wikipedia (unique)

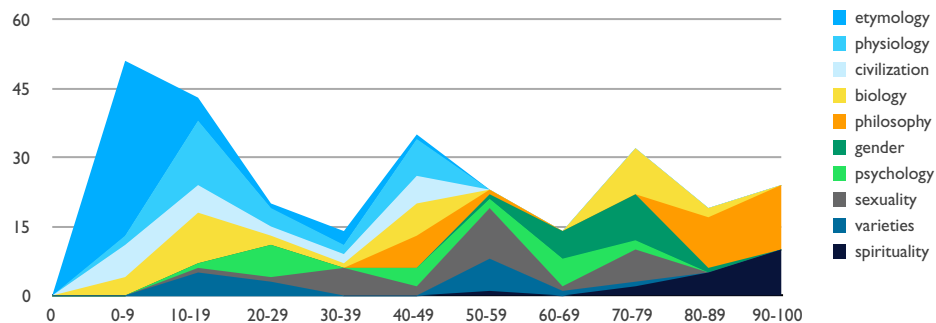
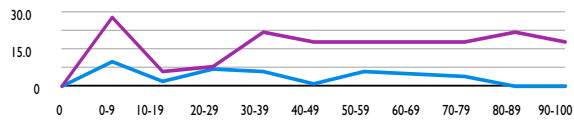


Figure 14

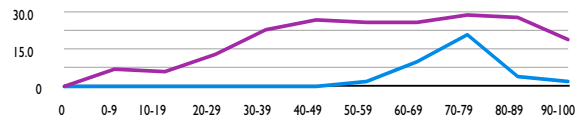
SEQUENCE: Comparison of article position between common words

— Encyclopedia Britannica — Wikipedia

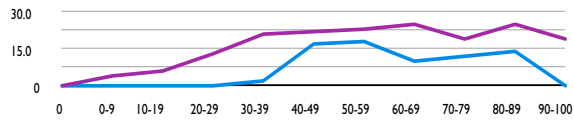
CLIMATE



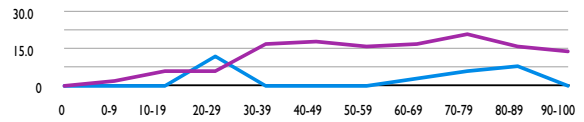
EDUCATION



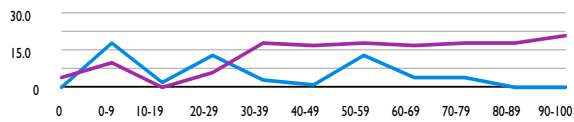
ECONOMY



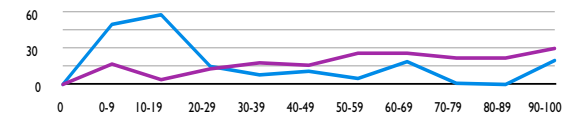
HEALTH



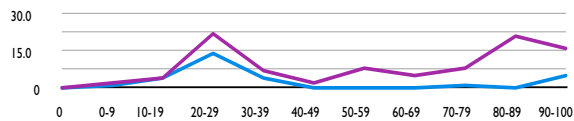
GEOGRAPHY



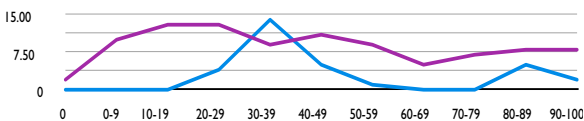
HISTORY



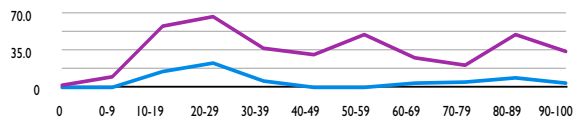
WORLD



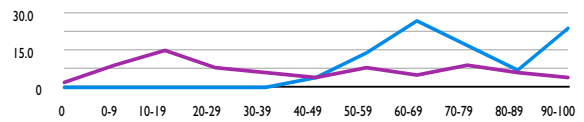
POLITICAL



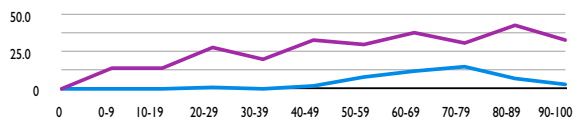
WAR



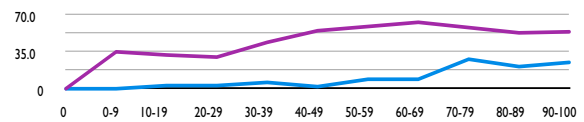
RELIGION



SOCIETY/SOCIAL



CULTURE/CULTURAL



GOVERNMENT

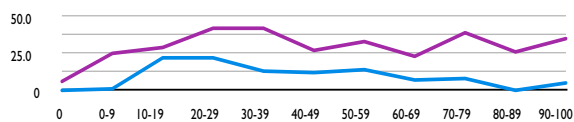
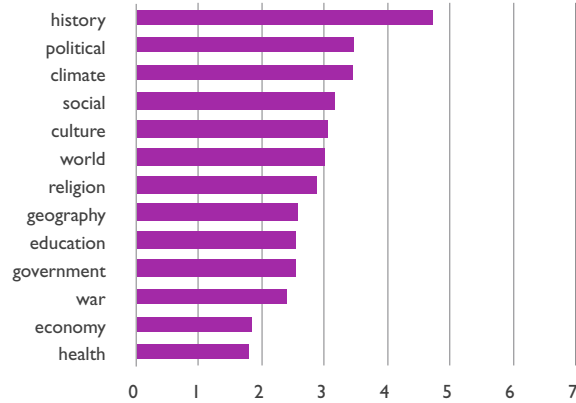


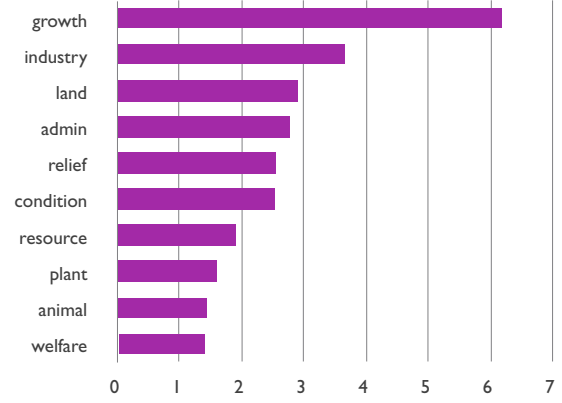
Figure 15
COMMAND: Ratio of heads:subheads

Encyclopedia Britannica

EB COMMON

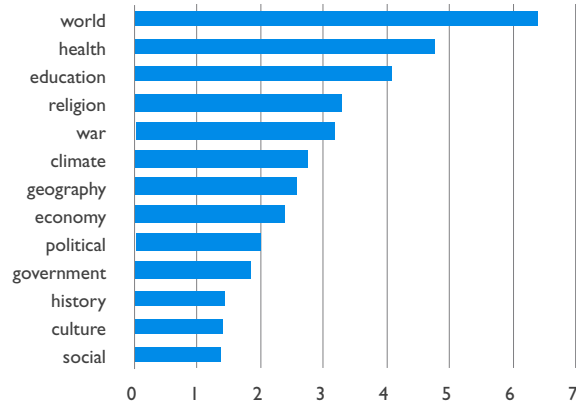


EB UNIQUE



Wikipedia

WK COMMON



WK UNIQUE

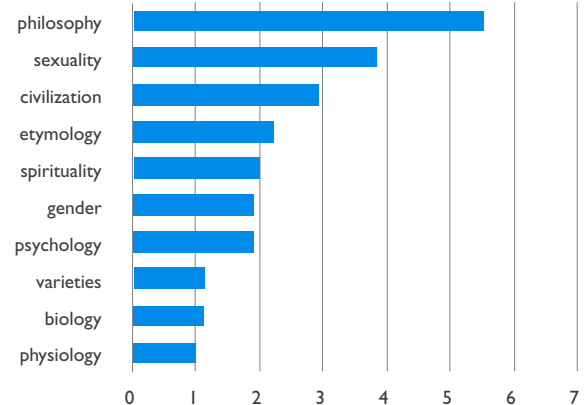


Figure 16
CONSISTENCY: Number of articles word is present in

— Encyclopedia Britannica — Wikipedia

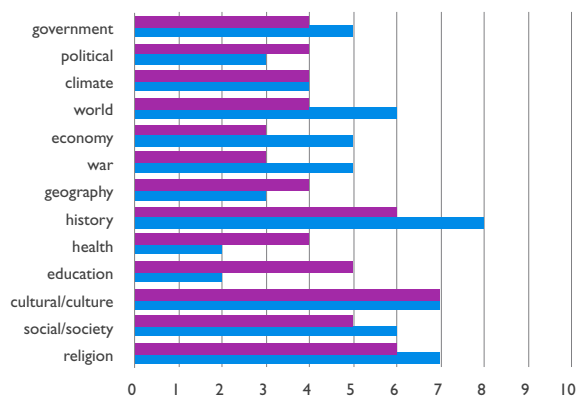


Figure 17
LONGEVITY: Number of editions word is present in

— Encyclopedia Britannica — Wikipedia

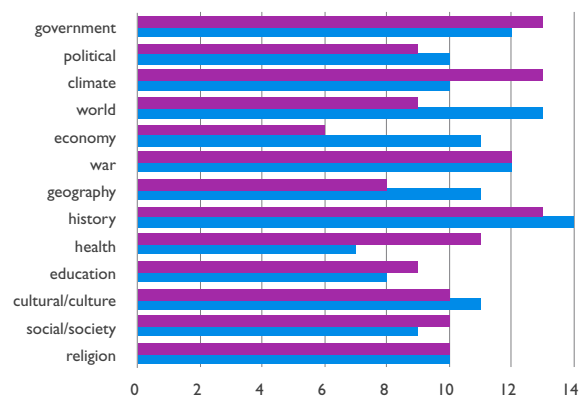


Table 15

CONSISTENCY: Number of articles word is present in

	AME	ATH	BLK	CDN	CNN	CAT	DEM	HUM	LON	OCE	SCORE
Britannica Common											
government	1	0	0	1	0	0	1	0	1	0	4
political	1	0	0	1	0	0	1	1	0	0	4
climate	1	0	0	1	0	0	0	0	1	1	4
world	1	0	0	1	0	0	0	0	1	1	4
economy	1	0	0	1	0	0	0	0	1	0	3
war	1	0	0	1	0	0	0	0	1	0	3
geography	1	0	0	1	0	0	0	0	1	1	4
history	1	0	1	1	0	1	0	0	1	1	6
health	1	0	0	1	0	1	0	0	1	0	4
education	1	0	0	1	0	0	1	1	1	0	5
cultural / culture	1	1	0	1	0	1	1	1	1	0	7
social / society	1	0	0	1	0	0	1	1	1	0	5
religion	1	1	0	1	0	1	0	1	1	0	6
Wikipedia Common											
government	1	0	0	1	0	0	1	1	1	0	5
political	1	0	0	1	0	0	1	0	0	0	3
climate	1	0	0	1	0	0	0	0	1	1	4
world	1	1	0	1	1	0	1	1	0	0	6
economy	1	0	1	1	0	0	0	0	1	1	5
war	1	1	0	1	0	0	1	1	0	0	5
geography	1	0	0	1	0	0	0	0	1	0	3
history	1	1	0	1	1	1	1	1	1	0	8
health	1	0	0	0	0	1	0	0	0	0	2
education	1	0	0	0	0	0	0	0	1	0	2
cultural / culture	1	1	0	1	0	1	1	1	1	0	7
social / society	1	1	0	0	0	1	1	1	1	0	6
religion	1	1	1	1	0	0	1	1	1	0	7
Britannica Unique											
administration	1	0	0	1	0	0	0	0	1	0	3
land	1	0	1	1	0	0	0	0	1	1	5
conditions	1	0	0	1	0	0	0	0	1	1	4
growth	1	0	0	1	0	0	0	1	1	1	5
industry	1	0	0	1	0	0	0	0	1	0	3
relief	1	0	0	1	0	0	0	0	1	1	4
animal	1	0	0	1	0	0	0	0	1	0	3
welfare	1	0	0	1	0	0	0	0	1	0	3
plant	1	0	0	1	0	1	0	0	1	1	5
resources	1	0	0	1	0	0	1	0	0	1	4
Wikipedia Unique											
etymology	1	1	1	1	1	1	0	1	1	0	8
biology	0	0	0	0	0	0	0	1	0	1	2
philosophy	1	1	1	0	0	0	0	1	0	0	4
physiology	0	0	0	0	0	1	0	1	0	0	2
sexuality	0	0	1	0	0	0	0	1	0	0	2
civilization	0	0	0	0	0	0	0	1	0	0	1
gender	0	0	0	0	0	0	1	1	0	0	2
psychology	0	0	0	0	0	0	0	1	0	0	1
spirituality	0	1	0	0	0	0	0	1	0	0	2
varieties	0	0	0	0	0	1	1	0	0	0	2

Table 16

LONGEVITY: Number of editions word is present in

	A 1771	B 1778	C 1790	D 1798	E 1842	F 1853	G 1878	H 1910	I 1968	J 1974	K 1984	L 1990	M 2005	N 2011	SCORE
Britannica Common															
government	1	0	1	1	1	1	1	1	1	1	1	1	1	1	13
political	0	0	0	0	0	1	1	1	1	1	1	1	1	1	9
climate	0	1	1	1	1	1	1	1	1	1	1	1	1	1	13
world	0	1	1	1	0	0	0	0	1	1	1	1	1	1	9
economy	0	0	0	0	0	0	0	0	1	1	1	1	1	1	6
war	0	1	1	1	0	1	1	1	1	1	1	1	1	1	12
geography	0	0	0	0	1	1	1	1	1	0	0	1	1	1	8
history	0	1	1	1	1	1	1	1	1	1	1	1	1	1	13
health	0	0	1	1	0	1	1	1	1	1	1	1	1	1	11
education	0	0	0	0	0	1	1	1	1	1	1	1	1	1	9
cultural / culture	0	0	0	0	1	1	1	1	1	1	1	1	1	1	10
social / society	0	0	0	0	1	1	1	1	1	1	1	1	1	1	10
religion	0	0	0	0	1	1	1	1	1	1	1	1	1	1	10
Wikipedia Common															
government	0	0	1	1	1	1	1	1	1	1	1	1	1	1	12
political	0	0	0	0	1	1	1	1	1	1	1	1	1	1	10
climate	0	0	0	0	1	1	1	1	1	1	1	1	1	1	10
world	1	1	1	1	0	1	1	1	1	1	1	1	1	1	13
economy	0	0	0	1	1	1	1	1	1	1	1	1	1	1	11
war	0	1	1	1	1	0	1	1	1	1	1	1	1	1	12
geography	0	0	0	1	1	1	1	1	1	1	1	1	1	1	11
history	1	1	1	1	1	1	1	1	1	1	1	1	1	1	14
health	0	0	0	0	0	0	0	1	1	1	1	1	1	1	7
education	0	0	0	0	0	0	1	1	1	1	1	1	1	1	8
cultural / culture	0	0	0	1	1	1	1	1	1	1	1	1	1	1	11
social / society	0	0	0	0	0	1	1	1	1	1	1	1	1	1	9
religion	0	0	0	0	1	1	1	1	1	1	1	1	1	1	10
Britannica Unique															
administration	0	0	1	1	0	1	1	1	1	1	1	1	1	1	11
land	0	0	1	1	1	1	1	1	1	1	1	1	1	1	12
conditions	0	1	1	1	0	0	1	1	1	1	1	1	1	1	11
growth	0	0	1	1	1	1	1	1	1	1	1	1	1	1	12
industry	0	0	1	1	0	0	1	0	1	1	1	1	1	1	9
relief	0	0	1	1	1	0	0	1	1	1	1	1	1	1	10
animal	0	1	1	1	1	1	1	0	1	1	1	1	1	1	12
welfare	0	0	0	0	0	0	0	0	1	1	1	1	1	1	6
plant	0	0	0	0	1	0	1	1	1	0	0	1	1	1	7
resources	0	0	0	0	0	0	1	0	0	1	1	1	1	1	6
Wikipedia Unique															
etymology	0	0	0	0	1	1	1	1	1	1	1	1	1	1	10
biology	0	0	0	0	0	1	1	1	1	1	1	1	1	1	9
philosophy	0	0	0	0	1	0	1	1	1	1	1	1	1	1	9
physiology	0	0	0	0	0	0	1	1	1	1	1	1	1	1	8
sexuality	0	0	0	0	0	0	1	1	1	1	1	1	1	1	8
civilization	0	0	0	0	0	0	0	1	1	1	1	1	1	1	7
gender	1	1	1	0	1	0	0	1	1	1	1	1	1	1	11
psychology	0	0	0	0	0	0	1	1	1	1	1	1	1	1	8
spirituality	0	0	0	0	0	1	1	1	1	1	1	1	1	1	9
varieties	0	0	0	0	1	1	1	1	1	1	1	1	1	1	10

Figure 18

Final scores for common and unique words

	Longevity	Consistency	Command	Rank	Sequence	Final Score
BRITANNICA COMMON WORDS						
1	climate	cultural/culture	geography	cultural/culture	cultural/culture	cultural/culture
2	government	history	history	war	war	war
3	history	religion	economy	government	government	government
4	war	social/society	cultural/culture	social/society	social/society	social/society
5	health	education	political	history	history	history
6	cultural/culture	climate	social/society	economy	economy	climate
7	religion	government	government	geography	geography	education
8	social/society	health	world	education	education	economy
9	education	political	religion	climate	climate	geography
10	political	world	climate	health	health	health
11	world	geography	war	world	world	political
12	geography	war	health	political	political	world
13	economy	economy	education	religion	religion	religion
WIKIPEDIA COMMON WORDS						
1	history	history	geography	history	history	history
2	world	cultural/culture	education	cultural/culture	government	government
3	government	religion	history	government	geography	cultural/culture
4	war	world	cultural/culture	religion	war	geography
5	cultural/culture	social/society	religion	economy	cultural/culture	war
6	economy	government	social/society	geography	economy	climate
7	geography	war	war	war	religion	religion
8	climate	economy	economy	social/society	climate	economy
9	political	climate	climate	climate	world	social/society
10	religion	geography	political	education	political	world
11	social/society	political	government	political	social/society	political
12	education	education	health	world	health	education
13	health	health	world	health	education	health
BRITANNICA UNIQUE WORDS						
1	animal	growth	conditions	land	land	land
2	growth	land	plant	administration	administration	administration
3	land	plant	administration	conditions	growth	growth
4	administration	conditions	growth	growth	animal	conditions
5	conditions	relief	land	animal	conditions	animal
6	relief	resources	resources	relief	industry	industry
7	industry	animal	industry	industry	relief	plant
8	plant	administration	relief	welfare	welfare	relief
9	resources	industry	welfare	plant	plant	welfare
10	welfare	welfare	animal	resources	resources	resources
WIKIPEDIA UNIQUE WORDS						
1	gender	etymology	varieties	etymology	etymology	etymology
2	etymology	philosophy	physiology	biology	physiology	biology
3	varieties	gender	spirituality	physiology	biology	sexuality
4	biology	varieties	civilization	philosophy	civilization	physiology
5	philosophy	biology	philosophy	sexuality	sexuality	psychology
6	spirituality	spirituality	biology	psychology	psychology	civilization
7	physiology	physiology	psychology	civilization	varieties	varieties
8	psychology	sexuality	gender	spirituality	philosophy	philosophy
9	sexuality	psychology	etymology	gender	gender	gender
10	civilization	civilization	sexuality	varieties	spirituality	spirituality

Figure 19
Comparison of most influential words

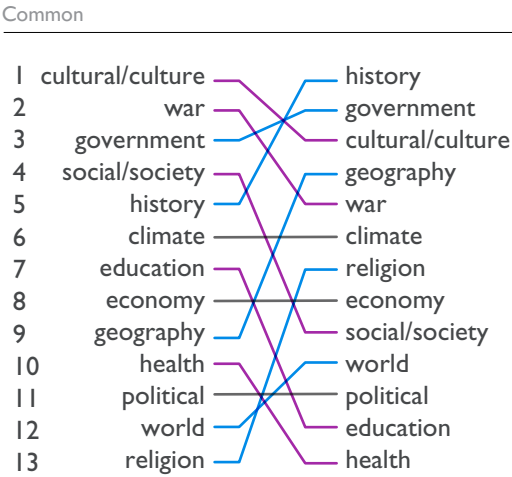


Figure 20
Comparison of most influential words

