Sum of all knowledge: Wikipedia and the encyclopedic urge

Salor, F.E.

2. Imagining and Building a New Encyclopedia

> It is quite natural that such developments should seem like marvels and that initially, deeming them to be impossible, the mind should reject any pursuit of them. But, according to a common slogan, do we not live in a time in which yesterday's utopia is today's dream and tomorrow's reality?
>
> Paul Otlet
> *Microphotographic Book*, 1906

2.1 Introduction

The modern encyclopedia, as it existed in the beginning of the 20th century, was the culmination of a continuous evolution that can be traced back to its Enlightenment origins, which in turn was the transformation of a much older ideal. The previous chapter traced the evolution of the encyclopedic form in line with its surrounding cultural, economic and technical environments. Encyclopedias occasionally fell behind the curve, with the continuing use of Medieval works during Renaissance, while sometimes they were the pioneers of new ideas that transformed the rest of the society, famously in the case of the *Encyclopédie*. At the end of the 19th century, the most notable encyclopedias settled into a matured form and focused their editorial efforts on increasing their breadth and depth, and this is exemplified by acclaimed 9th edition of *Britannica*.

Along with their intellectual ambitions, encyclopedias also become significant business ventures, involving considerable capital investment and marketing budgets. While the struggle to combine the right amount of business acumen with excellence in scholarship has worked very effectively in refining the model, a radical rethinking of the encyclopedic ideal in its entirety would be unlikely to emerge from the offices of any large and established encyclopedia. Still, the beginning of the 20th century witnessed some of the most ambitious ideas concerning the future of encyclopedias from some of the most original and/or seminal thinkers of the period. Others, building upon such ideals and inspired by rapidly advancing

technology, expanded their vision of a fundamentally new encyclopedia, among which *Wikipedia*'s roots can be traced.

This chapter aims to bridge the gap between mature modern encyclopedias as they were established at the end of the 19th century, and *Wikipedia*'s emergence in the beginning of the 21st century. This survey of the 20th century aims to cover two broad and intertwined themes, beginning with the earliest attempts to re-imagine encyclopedic thinking, and follows the works of theorists and science fiction authors. Some of these writers have laid down the foundations of our current understanding of digital technology while others have speculated on a future where culture and knowledge have transcended their existing boundaries. The second theme will focus on the work of pioneering computer scientists and engineers who, throughout the 20th century, built digital computers and the Internet, and directly contributed to the realization of some of the visions outlined in the first section. The central aim of this chapter is therefore to identify the blend of technical innovation and intellectual vision that went into the creation of *Wikipedia* through a discussion of these two movements and their interplay throughout the 20th century.

2.2 Challenging Print

The massive increase in the number of printed books resulting from the proliferation of the printing press across Europe and the intellectual activity it fostered was illustrated in the previous chapter. The modern encyclopedia has come to incorporate a claim to address a perceived need to summarize the expanding body of knowledge that the development of print made possible. This trend can be traced back to Chambers' claim that his *Cyclopaedia* was the only book a layperson would ever need and the editorial struggle of Chambers' successors grew along with the accelerating advance of modern sciences. The result was often an expansion in the size of the encyclopedia coupled with editorial efforts to try sustaining a level of cohesion and accessibility across the whole set. While this was a particularly trying challenge for encyclopedia editors, the issue of increasing multitudes of books and expanding knowledge was relevant for a diversity of professionals such as librarians, bibliographers and lexicographers.

Belgian bibliographer Paul Otlet (1868-1944) was among the first to address the issue of organizing knowledge with a sensibility that would become increasingly familiar throughout the 20th century. A lawyer by profession, Otlet harbored a lifelong fascination with collecting and organizing knowledge. Rayward outlines Otlet's primarily concerns with a series of questions (2):

i. How best was order to be introduced into this proliferating, disorderly mass in such a way that progress in the world of learning could continue efficiently and effectively?
ii. How could rapid developments in all areas of knowledge, so characteristic of modern period, be mobilized for the benefit of society?
iii. How could the international flow of information, then obstructed (as it still is) by political, social and linguistic barriers on the one hand, and by cumbersome, unresponsive systems of publication, distribution and bibliographic processing on the other, become more open and more effective?
iv. How could accurate, up-to-date, "integrated" information tailored specifically and exactly to particular needs be derived from this mass, re-worked to form ensuring immediate and optimal usefulness, and made available without hindrance or delay, whatever such potentially infinite, unpredictable needs might be.

In his "Something about Bibliography", originally published in 1891, Otlet attempted to address some of these questions with a radically new vision for the organization of information. According to Otlet, it is an almost impossible task to organize all the publications in print and make them available in an accessible form. To address this problem, Otlet argued, "various parts of any book can be reduced to different elements [...] and a re-arrangement of contents not along the lines of the special plan of a particular book, but according to the genus and species appropriate to each element does not make for an loss of substance" (17). Underlying this ambitious proposal is Otlet's belief that "the external make-up of a book, its format and the personality of its author are unimportant provided that its substance, its sources of information and its conclusions are preserved and can be made an integral part at the organization of knowledge" (ibid). Otlet envisioned the resulting compendium as an impersonal, collectively produced "artificial brain by means of cards containing actual information or simply notes of references" (ibid.). Otlet also noted that his proposed system of cards would have the added benefit of not being bound with alphabetical organization (18).

With such a scheme, Otlet aimed to create a compendium of knowledge where "any one person should be able to rapidly find out, not all about a subject, but all about the branches into which it has been divided and about the work already done relative to those branches" (19). Otlet's desire to transcend the limitations imposed by print media and the other physical attributes of books are apparent in his proposals as is his willingness to embrace new technologies that might further his goal. Most practical applications of Otlet's thinking found their manifestations with the help of index cards, although he was very interested in the possibilities offered by microphotography and other technical innovations.

Otlet's ideas culminated in the creation of the *Mundaneum* in 1910, co-founded with his lawyer colleague Henri La Fontaine and financially backed by the Belgian government. Based on a sanctioned adaptation of Dewey's decimal system created by Otlet, called *Universal Decimal Classification* (UDC), the *Mundaneum* aimed to collect and organize the world's information on a modular system of index cards. By reducing all the separate publications to their essence and combining them under a comprehensive system of classification, Otlet was striving to compile "the Source, the permanent Encyclopedia, the Summa [that] will replace chaos with a Cosmos" (Otlet, 1990,83). Based on his experiences with the *Mundaneum*, Otlet published "Transformations in the Bibliographical Apparatus of the Sciences" in 1918, in which he detailed his vision for the future and explained his system of categorization. Sadly, Otlet's ideas lost their popularity over time, and the Belgian government withdrew its funding from the *Mundaneum* which was later occupied during WWII.

While Otlet's proposed database is not strictly encyclopedic in nature, his thought process is a descendant of the goals shared by most encyclopedists throughout the ages, with a particularly positivist twist, characteristic of the early 20th century. Rayward acknowledges this mechanistic structure in Otlet's thinking and cites the influence of 19th century positivist thought on the formulation of his ideas, especially manifested in Otlet's strong emphasis on reducing each work down to its essential facts (6).[8] While his radical approach

[8] Originating from Auguste Comte's writings, *Positivism* posited itself as a strictly empiricist approach to the philosophy of science. Rejecting metaphysics entirely, the positivist approach argued for a scientific, fact-

to bibliography failed to come to full fruition, Otlet's work is still significant as he was among the first to propose a radical restructuring of bibliographies to greatly advance their effectiveness. He also proposed a new approach to encyclopedic knowledge that would require a reassessment of the priorities of compilers.

In his attempt to address central questions regarding knowledge and its organization, Otlet also pioneered the urge to transcend print through other technical innovations. His search for a better medium was to be continued throughout the 20th century, culminating in the hyperlinked, online articles of *Wikipedia*. Otlet's revolutionary ideas situate him as the first precursor of *Wikipedia* in the 20th century, and the degree to which he envisioned the model that was to be created in the 21st century is most certainly noteworthy. As Joseph Reagle argues, "[t]he Repertory was international, multilingual, collaborative and predicated on technological possibility, much like *Wikipedia*" (21).

Following Otlet, the next notable proposal for the recasting of encyclopedic conventions came from author H.G. Wells, who is mostly remembered for being one of the founding fathers of modern science fiction. In a lecture delivered at the Royal Institution of Great Britain in 1936, Wells acknowledged his interest in compiling outlines and summaries. After stating the "ineffectiveness of modern knowledge and trained and studied thought in contemporary affairs", Wells goes on to illustrate his ambitious proposal to amend this problem (Wells, 1938:3). Echoing Otlet, Wells proposed the foundation of a *World Encyclopedia*, which would serve as a social organ and bridge the gap between the immense knowledge being created around the world and the people in positions of decision-making who are oblivious to it (11).

> World Encyclopedia would be a row of volumes…[where] he would find in clear understandable language, and kept up to date, the ruling concepts of our social order, the outlines and main particulars in all fields of knowledge, an exact and reasonably detailed picture of our universe, a general history of the word, and if by any chance he wanted to pursue a question into its ultimate detail, a trustworthy and complete system of reference to primary sources of knowledge. In fields where wide varieties of method and opinion exited, he would find, not casual summaries of opinions, but very carefully chosen and correlated statements and arguments (13).

---

based, study of social issues. Since Comte's proposal to establish Sociology as a science of society, many critics have pointed out the incompatibility of human social issues with methods of natural science.

It is clear from the above description that Wells' *World Encyclopedia* aspires to be much more than a compilation. In fact, Wells outright declares the established conventions of compiling encyclopedias outdated and inadequate (ibid). Wells argues that the *World Encyclopedia* should not be merely an assembly of fact and statement, but should act as "an organ of adjustment and adjudication, a clearinghouse of misunderstanding" (16). While editors of different encyclopedias throughout history have been ambitious to various degrees with their claims to authority, Wells aspires to a whole new level relevancy when he argues that the newly created work should become "an undogmatic Bible to a world culture" (14).

While striving to provide universal access to the sum total of human knowledge, the authoritarian undertones of Wells' vision can hardly be dismissed as peripheral to the principal goal. He attributes great importance to presenting a definitive consensus on every topic "assembled with the approval of outstanding authorities in each subject" (14). Given access to such a resource, Wells believes that most disagreements plaguing the world could be made to disappear and his world brain could "hold the world together mentally" (ibid). Behind Wells's conviction lies his belief that most differences of opinion in the world are overrated and superficial (15). By giving everybody access to the same knowledge, verified by the proper authorities, Wells argues that the *World Brain* would "compel men to come to terms with one another" (16). While suggesting such an ambitious scheme, Wells does not address the questions of authority and validation that such an undertaking is bound to create, apart from reaffirming his preference for "order in place of vermin [and] a garden to a swamp" (ibid).

Although his suggestion that encyclopedic organization can spread like a nervous system places him quite close to the hyperlinked network of *Wikipedia*, in all other aspects, it would be safe to argue that Wells would have disapproved of an open encyclopedia that anyone can edit. Given all its shortcomings, it is easy to dismiss Wells's proposal as an early 20th century internationalist relic that claimed that English is the best language for such a project due to its linguistic superiority. When taken together with Otlet's ideals however, Wells' ideas clearly hint at an emerging trend.

As noted, the end of the 19th century saw the maturation of the encyclopedic ideal and, combined with the industrialization of the printing press, both the mechanical and intellectual processes of compiling and publishing encyclopedias have achieved their pinnacle of accomplishment. However, the 20th century witnessed the acceleration of the creation of new knowledge and the limitations inherent in the established methods of encyclopedia making, however refined they may have been, became increasingly evident. Otlet and Wells were among the first to realize that the only way out for encyclopedias was innovation. While their particular ideas were flawed and imperfect, and were still reliant on print, they vocalized some of the key themes that will be built on by their successors, such as universal access, constant maintenance and collaborative production. In addition, Wells may also be credited as the first science fiction author to think about encyclopedias, setting an example for many to come in the following decades.

The next section of this chapter will cover the propositions put forward by literary and scientific minds of the 20th century. However, to demonstrate the need to look at these sources that are usually understood as falling outside the trade, it is worthwhile to review a small sample of opinions on the future of encyclopedias from scholars and encyclopedists from the mid-20th century.

2.3 Maintaining the Modern Encyclopedia

The September 1962 issue of the journal *The American Behavioral Scientist* provides a selection of perspectives on the "uses of encyclopedias: past, present and future". Livio Stecchini's opening essay "On Encyclopedias in Time and Space" acknowledges the difficulty of creating a world encyclopedia, citing the near-impossibility of reaching a global agreement on most topics, and proposes to follow Pliny's example by starting with technical or scientific topics such as mathematics and physics (6). As I will discuss in greater detail below, such a technical bias has been observed in *Wikipedia* and while this has been criticized as a weakness, it also demonstrates the relative ease with which contributors are able to reach a consensus on such topics, validating Stecchini's insight almost 60 years later.

The following, and similarly themed article by Jacques Barzun focuses on the difficulty in addressing both scholars and newcomers to a given topic. Barzun considers this problem as a challenge for the editors of encyclopedias, yet never speculates on the possibility that such an arbitrary limitation can be overcome by moving beyond print. While the author acknowledges the challenge in achieving neutrality in encyclopedia entries, Barzun's proposal to define neutrality as "detachment and fair play with respect to definable issues" comes surprisingly close to *Wikipedia*'s central policy of *Neutral Point of View* (10).[9] Despite this principle of fair play, however, the author is quite certain as to which topics are, and will remain, worthy of deeper study. "Most small nations will remain small, and even if troublesome they may not be important from the point of view of civilization. And the older small nations will continue to warrant an interest that the new cannot claim; Switzerland is worth more study than Liberia" (10).[10] Later in his essay, Barzun further elucidates his approach, which is reminiscent of Wells's *World Brain*:

> A world encyclopedia in the strict sense will not be possible until the establishment of a world civilization. For the present, the phrase must continue to mean the circle of knowledge at command in the civilization now dominant, the western civilization which even its enemies and victims by the simple act of trying to become part of it, acknowledge as superior. (10)

The persistence of this particular brand of internationalism that was so evident in Wells, is noteworthy in an academic journal in 1962 and serves as a useful benchmark to frame some of the discussion concerning assumptions regarding knowledge contained in *Wikipedia*. Also, given his view on knowledge, it is not surprising that Barzun considers "letting 'small' or 'unimportant' articles be written by anybody who happens to be at hand" as "disastrous" (14).

In a very interesting contrast, the next article written by the editor in chief of the *Encyclopedia Britannica*, Harry Ashmore, readily accepts the historicity of the current editorial policy and the physical structure of the Britannica. While highlighting the role of

[9] Neutral Point of View (NPOV) is one of *Wikipedia*'s core content policies and defines the overall tone of *Wikipedia* content. NPOV, as a policy and as a broader approach to encyclopedic knowledge, has a central importance in understanding *Wikipedia* and will be discussed in detail in the next chapter.

[10] *A Disclaimer:* While such arguments are unthinkable in a contemporary academic journal and are certainly discomforting to quote in my dissertation, I still think it is illustrative of the struggle that the encyclopedia faced during the 20th Century.

Britannica as an instrument of popular education, the evolution of which is covered in the first chapter, Ashmore argues that in its current state, the Britannica represents the pinnacle of the form's evolution. However, in the concluding section of his brief article, Ashmore claims that the world of encyclopedias is on the brink of a breakthrough, one that made him feel "the chill breath of obsolescence on my own encyclopedic work" (18). Here he is writing about digital technology and the adoption of computers for organizing and accessing knowledge. Amidst the scholars who perceive the intellectual achievements of their time as universal and lasting truth, Ashmore's pragmatism and openness to the future resonates with some of the other visionaries that will be covered in this chapter.

2.4 Science, Fiction and Encyclopedia

Although encyclopedias are among the recurring objects that attract the attention of science fiction authors, one literary giant of the 20th century preceded all SF authors to the task of dismantling meaning attributed to print and its cultural attributes. Jorge Luis Borges (1899 – 1986), director of the Argentine National Library from 1955 to 1973, and one of the most acclaimed writers of 20th century, has reflected upon the nature of the book and the modern literate mind in a number of his stories. In "Garden of Forking Paths", Borges famously questioned the perceived linearity of time and cause-effect relationships people attribute to their actions. Replacing the singular passage of time with a labyrinth of forking paths that then fold into themselves, Borges's protagonist is conscious of the consequences resulting from each action of choice, regardless of how simple. Choosing an alternative is also rejecting all other possible universes in which such possibilities may exist. Borges visualizes an infinite number of universes differentiated by a single choice with varying ramifications, resulting not in linear time but rather in an "infinite series of times, a growing, dizzying web of divergent, convergent, and parallel times" (Borges, 85). According to Nick Montfort, the concept described in Borges's 1941 story is nothing short of a hypertext novel (29).

Similarly prescient is Borges's idea of an infinite library, containing every possible variation of 22 letters, bound in standard volumes, and inhabited by a bewildered group of

librarians who are born and die within it. By definition, The Library of Babel contains not only every book that has ever been written, but also all the books that can be written and their right and wrong translations of each other along, with erroneous collections and presumably a correct index. With such a dizzying concept, Borges gives his librarians what the encyclopedists of all ages desired: total knowledge. Borges then goes on to speculate on the possible implications of achieving such a collection. "When it was announced that the Library contained all books, the first reaction was unbounded joy. All men felt themselves the possessors of an intact and secret treasure. There was no [...] problem, whose eloquent solution did not exist-somewhere in some hexagon" of the library (69). But the true implications of their finding occurs to the librarians' shortly thereafter and their hopes of mastering all knowledge quickly turns into frustration since, for every book, every sentence and word that makes sense in the library's endless collection, there are an infinite number that do not. While presenting a fresh and critical look at one of the primeval urges of humanity, namely collecting and organizing knowledge, Borges's writing also questions the interaction between the limited, yet ever-curious, human mind and its relationship with information technologies, more specifically, language and print.

While Borges was trailblazing in the field of literature, engineers on the forefront of technological innovation, both during and after the Second World War, were imagining the next era of knowledge creation and storage. Beginning with observations similar to those made by Otlet and Borges, namely the apparent limitations of print, they went on to predict the computer revolution that would reshape the rest of the century. Following his central role in the Manhattan Project, the building of the atomic bomb, Vannevar Bush was keenly aware of the need to provide the next major goalpost for the nation's scientific community (Wardrip-Fruin, 35). In his 1945 *As We May Think*, Bush imagined the future of information storage, organization and retrieval based on his experience with early computers during the war effort. Echoing Otlet, Bush reiterates the inadequacy of established methods of transmission and storage due to the increasing specialization among experts and the growing mountain of data (37). His main issue with the established order of print-based information technologies is their incompatibility with the way the human mind works (44).

Bush is aware of the artificial linearity imposed by any index or categorization, where any object can only occupy one sub-category, and any attempt to introduce a more multiple system of categorization is not worth the extra labor required. In contrast, Bush states that the human mind operates on a principal of association, combined with random and instant access to each bit of information. To address this discrepancy, he acknowledges the potential of microphotography, especially for compressing vast amounts of information into tiny spaces. At the same time Bush was aware of the oncoming electronic computer revolution, and speculated on possible future stages of technical evolution.

His proposed system, the *Memex*, is "a device in which an individual stores all his books, records, and communications, and which is mechanized so that it may be consulted with exceeding speed and flexibility. It is an enlarged intimate supplement to his memory" (45). While the mechanical details provided by Bush for his device seem hopelessly out of date given the current state of computers, his impulse to provide an immediate and compatible extension to human memory based on associative organizational patters is as fresh as ever. In a bid to remain convincing in his seemingly daring proposal, Bush projected the technology of his day into the future. While the speed with which computers have evolved and became integrated into our daily lives would have been beyond the imagination of any scientist or science-fiction author in the early 1950s, the way in which Bush describes a typical use of the Memex machine is prescient of a typical researcher's experience with *Wikipedia* today. Bush imagines a researcher, starting from a simple question, following a diversity of links to various sources and being able to "build a trail of his interests through the maze of materials available to him" (46).

The authors discussed above set the tone for the rest of the 20th century with their writings. The further development and increasing ubiquity of digital computers arguably defined the scene of technology for the rest of the century and many authors, mostly in science fiction, continuously speculated on a future where access to and storage of knowledge has transcended the limitations of print. In addition to literary authors and engineering pioneers, the possibilities of new and future media technologies started to attract attention from a number of scholars during the mid-20th century as well. Before moving on to the developments of the latter half of the 20th century, it would be beneficial

and fitting to discuss some of the earliest work done on what will, in later decades, be called media studies.

2.5 Understanding Media, New and Old

While a definitive origin to the academic study of media might be elusive, the work of Harold Innis is a reasonable point of departure. Following a distinguished career in economics, Innis turned his attention to the interaction of specific communication media in different civilizations throughout the ages. In his 1950 volume *Empire and Communications*, Innis tentatively writes, "[i]t has seemed to me that the subject of communication offers possibilities in that it occupies a crucial position in the organization and administration of government and in turn empires and of Western civilization" (1950, 3). The fundamental concepts of Innis' analysis are time and space and his central claim is the existence of a tendency within different forms of communication to favor either time or space. This theme was first introduced in *Empire and Communications* and later analyzed in much more depth in the *Bias of Communication* (1951), both of which were published posthumously.

Innis states that "a medium of communication has an important influence on the dissemination of knowledge over space and time and it becomes necessary to study its characteristics in order to appraise its influence in its cultural setting" (1951, 33). The characteristics of a given medium would have an effect on the established notions of knowledge within a culture and, over longer periods of time, the effects would render the culture resistant to other media with different characteristics. In such a case, a radical shift resulting from the proliferation of a new medium would result in a dramatic restructuring and, in extreme cases, the construction of a different civilization.

It is a logical conclusion from these premises that a stable society would be based on a balance between these two extremes, but such a state is ever-elusive and, even when temporarily achieved, it can be disrupted with the invention of a new medium that might tip the balance one way or the other. In the chapter titled "Plea for Time", Innis states that visual media, like print and photography, emphasized individualism and created instability

within societies with "catchphrases like democracy, freedom of press and freedom of speech", only to be countered by the ear-oriented medium of radio and the rise of nationalist sentiments it led to (80). Innis is critical of all the innovative possibilities that seem to spring from newly developing media when he states that the potential of greater realism also carries within it the greater possibilities of delusion (72).

In his treatment of time, Innis is concerned with what later theorists have referred to as *immediacy* and the destruction of a sense of time through new media, where individuals are led to live "in the moment and for the moment" (90). For Innis, this development is the banishment of all individual continuity (90). Yet, while the possible effects of various media technologies on individuals became one of the most prominent issues of media studies after Innis, his focus was on political structures and how they might be affected by such changes in communication technology.

During his studies on the history of empires and their dominant methods of communication, Innis was fascinated by empires' ability to sustain themselves over vast geographical spaces and long stretches of time. Upon reviewing the history of various empires and their epochs, Innis concluded that every empire had a distinct method of administration and communication embedded in its social and political structure. Each of these distinct combinations marked an epoch in history that favors, and even builds upon, certain types of media. These media in turn determine the bias of the civilization in question.

Innis argued that this bias in communication also affects the culture and character of the civilization in question. A time-biased society would emphasize the stability and longevity of its customs and culture, a preference that would be clearly evident with the chosen durable materials on which communications are based. Such a society would give great importance to its customs, which would impede individualism. Ancient Egypt serves as a perfect example for a time-biased society, where the most important matters were always carved in stone and all the aspects of the culture were focused towards infinite stretches of time.

In contrast, a space-biased civilization would prioritize the future and the present more than the past. They would also tend to lose the sense of community that time-biased

societies hold in such high regard. Innis argues that contemporary western society is biased towards space in an unparalleled scope thanks to electronic technologies.

This distinction between media of communication also gave Innis a tool with which to analyze the transition from an oral-based culture to a culture that is based on writing. Concerning oral traditions, Innis states that societies structured by their tradition utilize the properties of oral communication and their strong religious sentiments to direct and enforce a body or cooperative community of individuals. Innis states that, in oral communication, all the senses--the eye, the ear and the brain--have to act together in supporting and competing roles, and these roles all have temporal characteristics. On the other hand, the detached and portable nature of print and the visual communication it entails, contributes to a civilization whose bias is rather towards space. The written word, the book as an object and its physical transportation are all oriented towards space, whereas speech is a time-dependent act that comes into existence and passes away with the passage of time.

Throughout his analysis, Innis emphasized issues that would ignite a discipline and outlined some of the key debates. His work has provided important insights for new generations, spearheaded by one of his colleagues, Marshall McLuhan, to whom the discipline of media studies arguably owes its fully-fledged existence. McLuhan is perhaps the most widely known and daring of this earliest generation of scholars who perceived that the relationship between media and society is critical enough to merit its own discipline in academia. A professor of English literature from the University of Toronto, McLuhan began his study of the interaction between media and culture with *The Mechanical Bride* (1951) in which he analyzed several advertising campaigns and their possible implications for the society. His following work, *The Gutenberg Galaxy* (1962) was far more influential despite the fact that McLuhan himself considered it as "a footnote to Harold Innis", and that he wrote it simply because someone had to.

McLuhan's primary objective in *The Gutenberg Galaxy* is to dissect the impact of the development and proliferation of the printing press on the western civilization. In his very detailed analysis covering fields of literature, social organization and thought, McLuhan provides ample evidence of the crucial importance of the printing press for the emergence of

western civilization. McLuhan attributes the rise of nationalism, scientific research and automation of production, which in turn partly led to the industrial revolution, to the proliferation of the printing press. According to McLuhan, the reason that the printing press was able to catalyze such sweeping changes throughout society was because of a shift it created in the ratios and priorities of sensory perception of among the literate portions of the society. *The Gutenberg Galaxy* also included some of McLuhan's most famous insights into the future, the most widely known of which is perhaps his prediction that the post-literate society of the future will form a Global Village. While McLuhan's preponderance for hyperbole limits the useful application of his work for a critical approach, it is still noteworthy that the *Gutenberg Galaxy* predates works like Eisenstein's by two decades and can be credited as being instrumental in initiating research interest in the significance of print.

McLuhan's following work, *Understanding Media: The Extensions of Man* published in 1964, was far more ambitious in its scope and aimed to investigate the interaction between culture and a wide variety of tools, from communication media to credit cards and electric lights. This work, and its catchy main argument, is responsible for casting McLuhan as the-medium-is-the-message-man, a legacy that lives on today.

After echoing Innis in stating that "every culture and every language has its favorite model of perception and knowledge that it is inclined to prescribe for everybody and everything", McLuhan argues that the mark of the time he writes in is an aversion to imposed patterns, and an ambition to "declare our beings totally" (6). McLuhan suggests that, in order to understand the effects of various media of communication, psychological and social consequences should be given the priority.

The rationale for this is McLuhan's famous, and much abused, statement that the medium is the message. What is meant by this statement is not always entirely understood, and indeed McLuhan himself often complained that no one had fully grasped what he meant. McLuhan's primary concern was to distinguish the content of a medium from the medium itself. In arguing that the media of communication shape and control to a great extent the way humans perceive the world, McLuhan was concerned with the media itself as a mode, or a perspective. The actual content, whatever end to which one chooses to apply

that medium, is of no importance to McLuhan, a point he demonstrates by arguing that the electric light is a pure medium and "whether the light is being used for brain surgery or night baseball is a matter of indifference" (8). Furthermore, focusing the analysis on content can be dangerous since the content can blind the viewer to the true nature and effects of the medium itself. As an example to this disconnect between medium and content, McLuhan argues that what historians generally consider to be the emergence of rationality during the Enlightenment was actually consequences of the proliferation of the printing press taking hold over the mind set of Europe.

In an attempt to better understand the differences brought about by literacy and possible future developments, McLuhan introduced a division between hot and cool media. A cool medium provides little information and leaves gaps to be filled by the receiver, and is therefore inclusive and participatory. Cool media favor a tribal social structure wherein tribe members actively participate in the production of information and meaning. Hot media on the other hand, provides much more information and do not require the same amount of participation from users. Hot media are mechanical, repetitive and in that sense they are easier to grasp. In effect, hot media work against tribal structures. According to McLuhan, electric media also render geographic locations less important, since they allow instant action and reaction by moving with the speed of light. As such, electric media entail such a speeding up that the process of fragmentation achieved by previous modes of communication is reversed and the society enters a phase of retribalization and decentralization.

As the title of his work implies, underlying this analysis is McLuhan's argument that every tool and every gadget ever invented can be taken as an amplification device for one of our bodily parts. Technologies of communication are no exception and following this understanding, the "content" of any medium is always another medium, each successive one extending the one that came before it. Writing was an extension of speech and the primary function of the printing press was to amplify and extend writing to a wider scope and to increase availability. Print was later amplified by the telegraph in a similar fashion. Among all the other prosthetic enhancements like cars that extend our feet, and glasses that extend our capacity to see, McLuhan reserves a special place for the latest developed electric media.

He argues that electric media extend our very nervous system and tap directly into our consciousness, realigning our sense ratios and perceptions. As a result, connecting with the network of electric media implies a transcendence of our local, physical beings, with the effect that, "in the electric age, we wear all mankind as our skin" (52).

Despite his cryptic writing style and the half century that has passed since he published his main body of work, McLuhan's impact on the analysis of media is immense. Media studies become a burgeoning field following his widely publicized work and some of his writing is still inspirational for his daring attitude if not for his specific ideas.

2.6 Encyclopedia Galactica vs. The Guide

In tandem with Bush's essay, the digital computer revolution was starting to emerge in research laboratories across the United States. While these advances would eventually define the future of encyclopedias, it would take two decades for computers to become feasible for mass adoption. During the mid-20th century, therefore, imagining a future digital encyclopedia was still in the domain of the most speculative of science-fiction authors.

Isaac Asimov, one of the 20th century's great science fiction authors, coined the term *Encyclopedia Galactica* as the repository of all knowledge of the known universe. In Asimov's *Foundation*, first published in 1951, his encyclopedists retreat to the remotest corner of the Milky Way, seemingly to compile the ultimate encyclopedia. The real purpose of Asimov's Encyclopedists (or *the Foundation* as referred by the title) is to prepare the galaxy for approaching chaos and to shorten the "years of barbarism" by preserving critical technical know-how. While this effort proves effective only to a limited extent throughout the multi-volume tale of *The Foundation*, Asimov's reverence for the Enlightenment ideal of the encyclopedia is apparent. His encyclopedists are a group of experts and scientists, untouchable by any government authority while their barren planet is strictly off-limits to visitors from outside.

The Foundation is consistent with the generally mechanistic and scientific approach found in most of Asimov's works and his imaginary encyclopedia is probably most noteworthy for the reactions it inspired from later generation of sci-fi authors. Yet, while

many Sci-fi authors have tried to imagine ways that new technologies might affect the workflow of compilers, few have gone as far as Stanislaw Lem in representing a completely novel means of producing an encyclopedia. In his 1984 fictional preface to *Vestrand's Extelopedia in 44 Magnetomes,* Lem provides the advertising material for an encyclopedia written solely on the predictions of the future produced by very powerful computers. The editors of the *Extelopedia* argue that since all encyclopedias are out of date by the time they are printed, the only way to keep up is to prepare an encyclopedia not based on current facts but on predicted future events. Lem's imaginary preface and sales pitch is an ingenious satire of the constant struggle involved in preparing an encyclopedia, offering an ultimate solution for the frustration embedded in preparing a modern encyclopedia that has plagued the encyclopedists since 18th century. In addition to offering a possible way out of this constant struggle, Lem's Extelopedia also offers an interesting perspective on the role of computers in the compilation process. Although his fictional preface gives no indication that he considers such an enterprise to be desirable, Lem is unique in placing computers in charge of the actual writing of the encyclopedia, rather than casting them as enablers of better organization and writing, as evidenced in most other sci-fi narratives on this topic.

Both Asimov's and Lem's imaginations have offered a bold future for encyclopedia compilers. But, while these works have been a source of inspiration for many, their authors fail to demonstrate an achievable goal for encyclopedists and computer engineers after which to strive. Arguably the most widely known and influential science-fiction author who hypothesized on forms of encyclopedias was Douglas Adams, chiefly because he succeeded in striking a balance between an idea that is both futuristic yet realizable. Adams's *The Hitchhiker's Guide to the Galaxy* started its life as a BBC radio show in 1978 and was published as a novel in 1979. Alluding to Asimov, Adams introduces the Hitchhiker's guide as follows:

> In many of the more relaxed civilizations on the Outer Eastern Rim of the Galaxy, *the Hitchhiker's Guide* has already supplanted the great *Encyclopedia Galactica* as the standard repository of all knowledge and wisdom, for though it has many omissions and contains much that is apocryphal, or at least wildly inaccurate, it scores over the older, more pedestrian work in two important respects.
> First, it is slightly cheaper; and secondly it has the words DON'T PANIC inscribed in large friendly letters on its cover. (Adams, 79, 2)

The Guide consists of pages loaded on an electronic screen, accessed via search commands. If desired, the electronic device can read the entries aloud as well. Its random access and flat hierarchy allows extensive linking among articles and, by following these links, users can carve out their own reading path. Also, the electronic guide features a permanent connection to a data network that automatically keeps all its content up to date.

The amount of similarities between Adams's imaginary all-encompassing guidebook and most people's perception and daily use of *Wikipedia* has not gone unnoticed (Boutin, 2005). These similarities however do not end with the usage of and access to the guide. Adams's description of the *Guide*'s editorial policy brings to mind an understanding of human motivations cherished and cultivated by founders of *Wikipedia* and other online collaborative projects, namely the belief that people would be willing to contribute to building things without monetary compensation if they believe in the project. In *Life, the Universe and Everything* (1982), Adams writes, "...the *Guide* will employ anybody who wants to walk straight in off the street and get ripped off..." (117). Adams later explains that, since the editorial staff is accustomed to taking enormous lunch breaks, "most of the actual work got done by any passing stranger who happened to wander into the empty offices of an afternoon and saw something worth doing" (118). While Adams deliberately formulated his *Guide* to be a parody of such works as Asimov's solemn *Encyclopedia Galactica*, it is noteworthy that the *Guide* is among the first examples of an encyclopedic work that delegates the implicit trust in an encyclopedia not on the expertise of its authors but on the process resulting from the, seemingly random, actions of people with unqualified expertise that are mainly driven by curiosity.

Since the articles are written by researchers in the field and uploaded to individual copies automatically, and the editorial work is done by walk-ins who just find the work interesting, the guide faces many challenges where reliability and accuracy are concerned. On this matter, the official policy of the guide is, when the guide is inaccurate, it must be at least *definitively* inaccurate and the reality is often inconsistent (Adams, 1980, 30). Such a remark can be interpreted as a parody of the all-important role Asimov cast for his encyclopedists, who are charged with directing the future of humanity through their

omniscience, however, throughout his life, Adams was an avid follower of technology and was considered an astute observer of its cultural impact, so his remark can also be interpreted as another highlighting of the random and serendipitous nature of the knowledge creation process implied by the *Guide.* Perhaps the most clear illustration of his influence is the fact that, he is commonly acknowledged as an inspiration for *Wikipedia* by *Wikipedia*ns. While Adams is mainly considered to be a humor writer, his somewhat lighthearted approach towards creating and sustaining a (reasonably) reliable encyclopedia manifest itself in the contemporary *Wikipedia* community in surprising ways, as I will discuss while analyzing the motivations and internal dynamics of the contributors in the next chapter.

Moreover, it is not surprising that Asimov would not specify the role of technology in his work written in the 1950's whereas Adams and Lem can be much more specific and elaborate in their descriptions written in the 1980's. In the intervening three decades, computers had evolved from experimental machines that occupied whole buildings and took days to program for a single operation, into relatively accessible pieces of hardware that most enthusiasts could afford to own and tinker with as they pleased. While the personal computer revolution that would transform many aspects of culture still lay ahead, the trend was clearly apparent.

In the period after the latter half of 1980, a number of very influential science fiction works were published. However, with the increasing ubiquity of personal computers and Internet access, the focus of most authors has increasingly turned towards the future of these tools, resulting in works that surpass the objective of this study with their scope. One notable example of the further development of science fiction in the age of personal computers and the Internet is William Gibson's *Neuromancer* (1984). Gibson envisioned a future world of immersive virtual reality and the fusion of human and computer elements and discussed the implications of such developments on conceptions of self and awareness. Neuromancer inspired other works that explore similar issues in the following decades, culminating in a sub-genre called cyberpunk. [11] Among the most highly regarded works

[11] For an in-depth discussion of *Neuromancer* and Cyberpunk movement in general, see *Science Fiction* by Adam Roberts, Routledge: London, 2006, 125-133.

belonging to this sub-genre was Neal Stephenson's *Snow Crash* (1992), which also ruminates on the implications of a convergence between the human mind and computer networks. Stephenson, in a later interview, described the cyberpunk movement as science-fiction authors' attempt to catch-up with the rapidly advancing personal computer revolution (Stephenson, 2004).

While sci-fi remains as fertile a ground as ever for speculating on the future of social and technological evolution, any work that succeeds the above mentioned authors would postdate the launch of *Wikipedia* and therefore fall outside the scope of this study. However, during the brief period where the ever future looking gazes of sciences fiction authors coincided with a period of transformation for encyclopedias, the insights provided by these speculative authors was invaluable.

In this section, I have reviewed the ideas of some of the prominent science fiction authors who have speculated on the future of encyclopedias during the course of the 20th century. While they are not directly responsible for the creation of the tools that enabled the formation of *Wikipedia*, the authors like Asimov and Adams discussed above can be credited for questioning the established norms regarding encyclopedias in a manner that is very accessible to the general public and therefore they can be credited as instrumental in raising awareness that encyclopedias can be different.

2.7 Building the Third Revolution

To this point in this chapter I have outlined the intellectual history of searching for a better encyclopedia throughout the 20th century. The combined influence of the authors whom I have discussed in shaping the expectations and goals of *Wikipedia* creators and editors has been considerable, however the realization of these ideas to any degree was always dependent on the advancements in technology, specifically in electronic computers. The following section of this chapter will cover some of the underlying ideals that motivated the creators of the personal computer and the Internet as these relate to building an encyclopedia. Far from presenting an exhaustive history of networked personal computing,

the aim of this section is to trace the development and highlight some of the inherent properties of these technologies that were later harnessed by the *Wikipedia* community.

The revolutionary change in creating, organizing and accessing knowledge triggered by the proliferation of the printing press across Europe was addressed in the previous chapter. However, the limitations of print itself began to become apparent long before it was outlined in the work of the 20th century authors discussed above. It is worthy of note that one of the earliest attempts to surpass limitations of print can be attributed to Leibniz, previously mentioned as the last of the *living encyclopedias*.

An archetypical polymath, Leibniz was responsible for the ducal library in Hanover while working on his prolific body of work that encompassed mathematics, physics and philosophy. Although his invention of one of the earliest known cataloguing systems for books would have made him relevant for the present thesis, his work in mathematics is much more important and has far reaching consequences. Given his wide-ranging interests, Leibniz was acutely aware of the inefficiency of spending time working out arithmetic calculations that were necessary for his work yet, which did not constitute its central point. Leibniz hoped to relegate the tedious task of routine calculations to a mechanical apparatus he designed that operated on cogs and wheels (Dyson, 32). While designing a machine to perform routine calculations was very forward thinking in its own right, Leibniz was not unaware of the further potential that such a machine could unlock. Along with his calculation machine, Leibniz also worked out the binary logic that underlies all digital computers today and which led to the realization that any problem that can be formulated through such a notation can be put through the computer. There is no doubt that Leibniz was envisioning a future computer that can create, organize and access knowledge, given the necessary parameters.[12]

Leibniz's ideas relating to computing machines and the binary logic that power them lay mostly dormant for more than a century until they were revived by the British engineer

[12] The details and the accompanying mathematical foundations of binary logic and its evolution throughout the centuries following Leibniz falls outside the scope of this study. George Dyson's *Darwin Among the Machines* (1998) provides an in-depth look at the origins of digital computers and their fascinating evolution.

Charles Babbage for similar purposes. Like Leibniz, Babbage was motivated primarily by a desire to remove human error from routine calculations while saving time. His Difference Engine and Analytical Engine were mechanical prototypes of programmable computers of 20th century and Dyson argues that Babbage may be considered as the link that connects Leibniz's vision to the microprocessor driven computers of today (38). Like the computing machines of Leibniz, Babbage's difference engine of Babbage was prohibitively expensive and impractical given the contemporary technology, nevertheless, as Dyson states, Babbage's unrealized dream is the closest anyone got to creating the computer before electronic circuits rendered them viable.

According to Dyson, Babbage's pursuit of a precise and rapid computing machine was driven by his belief that computers would greatly enhance the process of creating and analyzing knowledge. Acquiring more complete knowledge, and faster, for Babbage meant a better understanding of God and his creations (Dyson, 42). The striking similarity of motivation between the 19th century Victorian engineer and medieval compilers of great encyclopedias may be considered a testament to the enduring quality of the encyclopedic urge.

While ideas that were put forward by Leibniz and Babbage pointed in the right direction, the true emergence of a programmable electronic computer coincides with the Second World War. Both Leibniz and Babbage dreamed of a machine that would take over the mundane and error-prone task of routine calculations, which constitute a significant part of any mathematical analysis or information processing effort. A similar routine calculation is also an integral part of cryptography, which is why the British government set up a facility at Bletchley Park, in the middle of Cambridge and Oxford, and staffed it with some of the most brilliant mathematicians and chess players of the era in order to break German codes, among which was the famous Enigma cipher. Leading this effort was mathematical prodigy Alan Turing, who was instrumental in the construction of *Turing Bombes* and the *Colossus*, the first programmable electronic computer in 1943. After the war, Turing published a paper titled "Computer Machinery and Intelligence" (1950) wherein he sketched out some of the fundamental ideas of computer science, such as the relationship between computers and language, and the features of a truly artificial

intelligence. Computers, especially in their networked form, operated with graphical user interfaces, did not become significantly important for encyclopedia makers until the end of the 20th century. Nevertheless, the ideas laid out by pioneers like Turing were instrumental in determining the direction of the encyclopedia's future evolution and laid the groundwork for the developments in computer science and engineering.

2.8 Hypertext

The development of computers continued rapidly after WWII, mostly in research laboratories in the United States. Under the direction of pioneers such as John von Neumann, the capabilities and speed of computers have grown exponentially. The idea of linking the existing research computers across the United States to form a rudimentary network was developed at a very early stage of the process, which later evolved into the Internet, which now covers most of the globe and connects billions of devices.

While the mechanical underpinnings of today's computers and the Internet were taking shape, Theodor Nelson was questioning their future impact on the way information is categorized and perceived. In his 1965 essay "A File System for the Complex, the Changing, and the Indeterminate", Nelson outlined the kind of structure he envisioned that computers might enable. In a similar vein, along with Bush, Nelson speculated on "the kinds of file structure required if we are to use the computer for personal files and as an adjunct to creativity" (134). Nelson is aware of the wide ranging potential of computers and he therefore proposes a "simple and generalized building-block structure, user oriented and wholly general-purpose" (ibid). Like Bush, who argued that computers would be better suited to organize information in ways similar to the human brain, Nelson claims that a computer can adapt its storage of knowledge to the particular needs of each user, enabling much greater usability. Nelson refers such a cluster of information that is interlinked and navigable by individual preference as *hypertext*. Nelson's ideas proved to be seminal for the modern Internet, which is built around the idea of hyperlinks and a dense and comprehensive linking policy is, in fact, one of *Wikipedia*'s core strengths today.

Many of the ideas proposed by Bush and Nelson became real-life possibilities in the 1980's when personal computers, with graphic user interfaces, became available at accessible prices to a much wider public. While the personal computer made desktop publishing possible, it also introduced the tools to finally transcend print and ushered in the widespread adoption of hypertext.[13] However, the interest in the non-linear, interlinked nature of hypertext has been strong in various branches of academia since the 1980's and numerous influential scholars explored the possibilities of hypertext literature in various forms.[14]

Given the technical advances and the increasingly widespread adoption of networked and multimedia capable computers, various publishers created digital versions of their encyclopedias. First published on optical storage media such as CD and DVD, and later presented as websites, the major differences between these incarnations of traditional encyclopedias and their print forms were the inclusion of richer multimedia content such as photos, videos and animation, as well as a search function that helped users to access individual articles. With the inclusion of personal bookmarks and note taking within the software, it could be argued that these digital versions of traditional encyclopedias took steps toward achieving Bush's idealized Memex, but their proprietary architecture prevented them from becoming a universal and personal knowledge repository.

Despite the constant advances made in the sphere of personal computers, isolated machines used by individuals remained limited in their impact on broader social issues. The proliferation of the Internet as a useable tool for ordinary PC owners marked a watershed moment in a revolution that was comparable in magnitude to the invention of the printing press centuries earlier. As previously noted, the idea of networking computers to exchange information is as old as the first computers that were developed in the 1950's. In the intervening years however, computers became much more familiar to a greater number of individuals who used them for far more diverse tasks than military applications such as

[13] *Desktop Publishing* refers to the introduction of graphic user interfaces (Windows, menus and buttons instead of the previously prevalent command-line and the mouse as a pointing device) with what-you-see-is-what-you-get text editing software. Combined, computers allowed text creators and editors to compose documents with all the flexibility afforded by computers while keeping track, in real-time, of how their documents would look in print.

[14] See; Aarseth, 1994. Bolter, 2001. Coover, 1992. Gunkel, 2004. Joyce, 1988, 2002. Maulthrop, 1991.

nuclear weapons research or national intelligence. Tim Berners-Lee, a researcher at CERN, Switzerland, adopted Bush and Nelson's ideas for hypertext in a multi computer network, and then made the technology freely available. It could be argued that the decision to make the web decentralized and freely adaptable is as important and central to its success as the underlying technology. Thanks to Berners-Lee, no government or organization can claim ownership or overarching control of the World Wide Web, and the open and adaptable structure of the network itself would allow it to simply route around any attempt to control it. Although partly based on technical considerations, such design decisions ingrained a specific mindset in many web-based initiatives, among which *Wikipedia* is a prime example. Principles such as openness, combined with the continuous decline of hardware costs led to an exponential growth rate of users. Indeed, while in 1995 there were 15 million users around the world, in ten years this number has jumped to more than one billion and, as of June 2010, it stands close to two billion. This rapid adoption brings to mind the proliferation of the printing press across Europe, covered earlier. While other forms of media that emerged during the 20th century, chiefly radio and TV, achieved similarly rapid mass adoption, the unique affordances of the Internet as a creative medium as opposed to a broadcast apparatus renders it a closer descendent of the so-called Gutenberg revolution than any other.

So far, I have detailed the emergence and mass adoption of the networked personal computer. Starting as a speculative idea of futurists in the 1950's, computers slowly emerged throughout research labs and continued to gain abilities for the rest of the century both by advances in hardware and most importantly, by being networked together. The potential of computers to transcend the boundaries set by physical media, from papyrus scrolls to mass produced volumes, was initially apparent. However, a wider social adoption of these technologies and their underpinning ideals was required for the emergence of actual projects, like *Wikipedia*, that challenge the established norms of print. While the development and adoption of the hardware was crucial in achieving this goal, an equally important component was the emergence of new methods of collaboration and related authorship practices that harness the possibilities of computers and the Internet. The next section of my analysis will trace the development of the open source software movement,

whose underpinning ideals and subsequent practices shaped collaborative knowledge production and authorship online, and therefore constitutes an integral component of understanding *Wikipedia*.

2.9 Open-Source and Collaborative Production

Although computer hardware became abundant and, with increasing competition, relatively low-cost, the operating systems and other software that ran on a majority of computers worldwide were developed and licensed by a relatively small number of companies. The open-source software movement that emerged as a reaction to closed software platforms became one of the most important business and social phenomena of the last decades and spearheaded an ideal of collaborative production on a previously unseen scale, thanks to the unique affordances of the Internet. This collaborative culture became one of the founding principles of *Wikipedia*, which in turn introduced the potential of collaborative production to a much larger audience.

The roots of the free and open-source software movements can be traced to the early 1980's and to Richard Stallman, a legendary programmer, at MIT. For Stallman, the practice of freely sharing work for the advancement of the community was an inherent aspect of software development, derived from the very nature of the code. Upon realizing that the commodification of software by corporations entailed arbitrary restrictions on code and therefore fragmented the efforts of programmers, he prepared *The GNU Manifesto*, outlining an alternative route. The central idea behind Stallman's *Manifesto* is that software should be free, not in a monetary sense, but rather in the ways in which people can interact with the software code in question. According to Stallman, users should have the freedom to use code in any way they desire, to be able to modify it and redistribute in a modified version without hindrance (Stallman, 1985). In order to propagate software written according to these principles, Stallman released his code on a license that allowed anyone to copy and, more importantly, to modify it. The only caveat for interested programmers is that they would have to release their modified version under the same conditions. This radical new licensing scheme came to be called the GNU General Public License (GPL).

While not embracing the entirety of Stallman's, increasingly radical, ideals regarding free software and its political and social implications, a large community of developers embraced the basic principle of sharing their work and started to form thriving communities of open-source software development. The most famous, and possibly daring, of these projects is widely known as *Linux*, a completely free and open-sourced operating system written from the ground up collaboratively by volunteers. What started as a hobby for devoted programmers around the world transformed into an economic phenomenon with implications for diverse industries and types of businesses and organizations. In addition to being one of the first online collaborative efforts on a very large scale, Linux, and other open-source projects, are also significant as they demonstrated that online collaboration on an open platform could be successful, both socially and financially, and served as public advocates for their model of creation. Because of their success, a project like *Wikipedia*, while not affiliated in any way, appeared more plausible.

However, the open-source software movement was not the only undertaking that harnessed the collaborative potential of the Internet. In fact, one of the earliest projects ever to be implemented over the newly established rudimentary computer networks was *Project Gutenberg*, which aims to provide free and easy access to all public domain books.[15] Founded in 1971 by Michael Hart who was looking for something useful to do with the computer time generously granted him at the University of Illinois, *Project Gutenberg* has grown over decades to encompass a library of 33,000 eBooks, available completely free and without restrictions (Hart, 1992).

The first text to be entered into the repository was the "Declaration of Independence", chosen for its briefness and manually typed by Hart himself. In the following years Hart continued entering texts into the repository with occasional help from a small number of volunteers. Starting with texts like the "Constitution of the United States", Hart then moved on to other works of literature that are in the public domain. According to Marie Lebert's 2008 article on the history of the project, Hart himself was responsible for typing in almost all of the texts during the first twenty years of the project's existence.

[15] www.gutenberg.org

Project Gutenberg had a collection of 10 books by 1989, however the mass adoption of personal computers and the World Wide Web throughout the 1990s would completely change the pace of growth and accessibility of the repository. One could argue that the project found its natural home online where electronic books can be created and circulated with immense speed and Hart's solitary hobby started to garner volunteers. According to Lebert, while only one book per month was added to the repository in 1991, by 1996 the number rose to 32 books per month. By the end of the year 2000, *Project Gutenberg* had amassed a library of 3000 books and by the end of 2001, up to 104 new books per month was being added (Lebert). Therefore given the age of *Project Gutenberg*, it is possible to track the proliferation of the Internet just by observing its growth rate.

While the public Internet made *Project Gutenberg* widely available, abandoning manual text entry in favor of scanners and Optical Character Recognition (OCR) software meant creating new eBooks can be undertaken much more rapidly. OCR works by scanning a printed page and determining the content by software algorithms and, although considerably faster than typing, it introduces a unique set of problems due to the non-uniform nature of fonts and degraded book pages. With the introduction of OCR, the task of the contributors shifted to copy-editing rather than typing individual pages. However, since most volunteers would be willing to work on only portions of any given text at a time, synchronizing their efforts became the main challenge of the project.

A very effective solution to the problem of coordinating volunteers working on a single project came from Charles Franks in October 2000 in the shape of Distributed Proofreaders, a website that allows volunteers to proofread one page at a time, thereby greatly facilitating many people to work on a single book simultaneously. The website asks volunteers to compare a scanned image with the text created by the OCR software and make the necessary corrections after which each text goes through a second proofreading by a more experienced editor. Books assembled this way are then compiled, formatted and indexed before being presented in the Project Gutenberg database. For users, the online book repository offers a single and easy to use text, however, thanks to the collaborative nature of the project, preparation of a single work ceases to become a daunting task for an individual.

Project Gutenberg, and Distributed Proofreaders, represents some of the core attributes of online collaborative projects. Creator Michael Hart and his early collaborators were driven only by their desire to contribute to the nascent online culture and they willingly donated their free time, typing and correcting texts without financial motives. While Project Gutenberg is not unique, or first, in harnessing voluntary participation of time and attention by dedicated volunteers, its nature as a purely textual enterprise and the editorial function it requires from its volunteers makes it a significant precursor to *Wikipedia*. Given two distinctly different projects like Linux and Project Gutenberg, a closer look at the characteristics of specific projects that are optimally suited to harnessing these kinds of social behaviors would be instrumental in understanding *Wikipedia*.

In his comprehensive analysis of networked production and distribution Yochai Benkler identifies two of these characteristics as modularity and granularity (100). Modularity refers to "the extent to which it [a project] can be broken down into smaller components, or modules, that can be independently produced before they are assembled into a whole" while granularity refers to "the size of the modules, in terms of the time and effort that an individual must invest in producing them" (ibid). Open-source software movement and *Project Gutenberg* are some of the best examples of the potential when these elements come together and provide a template with which to analyze *Wikipedia*. Benkler's analysis provides a simple, yet versatile framework to evaluate online collaboration efforts. In the case of *Project Gutenberg*, checking for errors on pages of a book produced by an OCR software can be done even on a word by word scale, providing excellent modularity. Meanwhile, by expecting only one page at a time from a single contributor both sets a reasonable expectation and sets a meaningful goal to achieve, motivating the volunteers. A similar pattern can be observed with open source software development, where the coding effort can be compartmentalized and assigned to individuals or groups. *Wikipedia* benefits from similar properties by offering to build an encyclopedia through an iterative process that allows experimentation and the ability to revert changes. Overall, Benkler's analysis proves apt and insightful in understanding under which conditions a collaborative online project can flourish.

Collaborative production has been the topic of many studies in the recent years, both academic and otherwise. Given the incredible power it demonstrates and the wide scope of its applications, only some of the discussion surrounding it is relevant for this study. However, its profound implications concerning authorship and authority are central for this thesis. Before moving on to the analysis of *Wikipedia*, I will discuss some of the important theoretical works that aim to historicize these developments from the viewpoint of media studies.

2.10 Transcending Print

Given the speed of its growth and the central position it increasingly came to occupy, much has been written that has aimed at explaining the impact of the Internet on societies. While a comprehensive review of this literature would fall outside the scope of this study, unique affordances of the Internet define key characteristics of *Wikipedia*. Therefore, highlighting some of the continuities and novelties of this new medium is essential to understanding the transformation of the encyclopedic urge to an open, online community. The six key areas identified by Eisenstein as being fundamentally affected by the proliferation of the printing press prove to be very helpful in providing a contrast.

Eisenstein argued that print greatly multiplied the output of texts while standardizing individual copies with respect to language, layout and fidelity to the author's intentions. These changes allowed faster and better communication among peers, both by forming of a common frame of reference, and by allowing more frequent and participatory discussions. Greater dissemination also made it possible for books to become ubiquitous objects rather than closely guarded relics. Likewise, with increased access, a heightened sense of ownership over texts and more intense communication among peers, the citation of proper resources in texts became standard practice. Eisenstein went on to argue that the cumulative effects of these changes were instrumental in the emergence of the Renaissance and the Scientific Revolution in Europe in the following centuries. Regarding the specific aspects highlighted by Eisenstein, it is reasonable to argue that the affordances of the Internet represents a similar effect regarding print, as print represented to handwritten

scrolls. However, as various media theorists argued, and as Eisenstein demonstrated concerning print, a mere amplification, or speeding up of process usually signifies a much fundamental shift in the social implications of media.

One of the first to investigate such effects of media, Harold Innis, in his study of empires and communication, built his analysis on the observation that a dramatic increase in the inherent speed of a dominant form of communication is not a trivial change. By increasing the potential speed of text-based communication to simultaneity, for example, the number of participants in a dialogue increases, theoretically, to infinity while rendering communication infinitely copiable with no marginal costs. Given these observations, the Internet arguably represents as fundamental a revolution as the printing press. That said however, it should be noted that these effects of the proliferation of the Internet are in no way limited to text-based communication and encompass all spheres of media.

In their 2000 volume titled *Remediation*, Bolter and Grusin offer a theoretical framework to understanding these changes in a wide spectrum of media with the concepts of *hypermediation* and *remediation*. Bolter and Grusin define hypermediacy as an attempt to reproduce the rich sensory experience of human perception by utilizing multiple instances mediation at the same time (34). The authors add that, most media aim to provide an immediate connection between the user and the object by striving to disappear, arguably the most notable example of which can be found in Renaissance paintings with linear perspectives. The artists strove to replicate what the eye would see by their advanced knowledge of perspective and geometry and therefore aimed to render the material nature of the painting invisible. In contrast, a hypermediated communication aims to achieve immediacy by making the user specifically aware of this mediated process (ibid.). Although instances of hypermediacy can be found throughout history, in medieval paintings for instance, the authors argue that the form reached its current zenith with the personal computer and the World Wide Web. A traditional graphical user interface (gui) found in almost all modern computers aims to render the computer more familiar and useable by providing users with familiar concepts like a desktop, folders, menus and icons.[16] While such

[16] The first personal computer to introduce a graphical user interface, and mouse, was the Apple Macintosh and it was introduced in 1984 with the advertising slogan "The computer for the rest of us".

a visual representation is much more easy to grasp as opposed to the previously dominant method of typing every single command by hand into a command line, the diverse and heterogeneous nature of the interface serves as a constant reminder of the interface itself, no matter how intuitive, and by definition of the authors qualify as hypermediate.

While the capacity of the Internet to combine all media and remediate them, to use Bolter and Grusin's terminology, is remarkable, the effects of the proliferation of the Internet on the production and organization of knowledge is of much more crucial importance when analyzing the transformation from the established 20th century encyclopedias to *Wikipedia*. However, in light of the ideas regarding the future of encyclopedias covered in this survey, it is clear that modern and ubiquitous Internet is a truly disruptive technology that is capable to transform encyclopedia in every aspect.