

## 2 The Pursuit of the Universal Encyclopedia

**sum of all human knowledge** *adj.*: Descriptive of the ultimate aim of Wikipedia. Regretfully, the vast majority of human knowledge is not, in actual fact, of interest to anyone, and the benefit of recording this collective total is dubious at best.

—WikiSpeak

In March 2000 Jimmy Wales, cofounder of Wikipedia and its Nupedia progenitor, sent his first message to the Nupedia email list: “My dream is that someday this encyclopedia will be available for just the cost of printing to schoolhouses across the world, including ‘3rd world’ countries that won’t be able to afford widespread internet access for years. How many African villages can afford a set of Britannicas? I suppose not many. . . .”<sup>1</sup> In this statement one can find a particular type of Enlightenment aspiration: a *universal* encyclopedic vision of increased information access and goodwill. Richard Schwab, a scholar of the Enlightenment and *Encyclopédie*, wrote that at that time thinkers were coming to recognize that “cumulatively they were a force in the world” and possessed “a new solidarity and power to advance human affairs.”<sup>2</sup> For example, Denis Diderot (1713–1784), editor of the *Encyclopédie*, wrote that it was developed by a society of men working separately on their respective tasks “but all bound together solely by their zeal for the best interests of the human race and the feeling of mutual good will.” The aim of this effort was “to collect all the knowledge that now lies scattered over the face of the earth, to make known its general structure to men among whom we live, and to transmit it to those who will come after us.”<sup>3</sup> Historian Richard Yeo describes the universalistic principles underlying so much of the Enlightenment culture as “the ideal of transportable knowledge, the communication of ideas across national and confessional boundaries; the ability of individuals, where ever they lived,

whatever their social status, to participate in a universal conversation. . . . These [reference] works offered the possibility of a reliable codification of knowledge by seeking to record any consensus, and by fixing the meaning of terms." As an example of unbridled optimism, Yeo cites the French social thinker and encyclopedist Henri Saint-Simon's (1760–1825) forecast that lasting peace could be achieved between France and England within a year of jointly undertaking work on a "New Encyclopedia."<sup>4</sup>

Granted, one can poke fun at the pretense of summarizing "all human knowledge," as is done in the WikiSpeak definition at the start of this chapter. Creating a world encyclopedia, much less world peace, is a difficult task and the trivia found on Wikipedia is a source of delight to some and derision to others. Nonetheless, the coupling of increased information access with human accord is a long-held dream. While its advocates are sometimes overly exuberant they can also be pragmatic, as seen in Wales's 2004 "Letter from the Founder":

Our mission is to give freely the sum of the world's knowledge to every single person on the planet in the language of their choice, under a free license, so that they can modify, adapt, reuse, or redistribute it, at will. And, by "every single person on the planet," I mean exactly that, so we have to remember that much of our target audience is not yet able to access the Internet reliably, if at all. . . . Our community already comes from a huge variety of backgrounds, and over time the variety will only increase. The only way we can coordinate our efforts in an efficient manner to achieve the goals we have set for ourselves, is to love our work and to love each other, even when we disagree. Mutual respect and a reasonable approach to disagreement are essential . . . on this incredible ridiculous crazy fun project to change the world.

None of us is perfect in these matters; such is the human condition. But each of us can try each day, in our editing, in our mailing list posts, in our irc [Internet Relay] chats, and in our private emails, to reach for a higher standard than the Internet usually encourages, a standard of rational benevolence and love.<sup>5</sup>

Furthermore, whereas Wales conceived of his encyclopedia reaching those without access to the Internet, technology is central to the modern version of the vision.<sup>6</sup> Technology is expected to facilitate a radically accessible resource that bridges the distance between people. As recounted in Tom Standage's history of the telegraph (i.e., the "Victorian Internet"), the "rapid distribution of news was thought to promote universal peace, truthfulness, and mutual understanding."<sup>7</sup> H. G. Wells felt that "Encyclopaedic enterprise has not kept pace with material progress" but when the

“modern facilities of transport, radio, [and] photographic reproduction” were embraced the creation of a permanent world encyclopedia would be “a way to world peace”: “Quietly and sanely this new encyclopaedia will, not so much overcome these archaic discords, as deprive them, steadily but imperceptibly, of their present reality.”<sup>8</sup> One can even see the universal vision in a different sort of technology altogether: the airplane. Joseph Corn, in *The Winged Gospel*, tells of high aeronautical expectations. “Air Globes,” representations of the earth and its cities without political or geographical boundaries, were deployed in the classroom to tangibly symbolize “the new world which Americans believed the airplane was about to create, a world of peace where national boundaries and topographical features were no longer pertinent.”<sup>9</sup>

This technological inspiration and aspiration for global accord is quite in keeping with the heritage of Wikipedia described in this chapter—and globes are a recurrent motif in this work. And while one can draw parallels between Wikipedia, the Enlightenment, and even aeronautics or ancient encyclopedic efforts, I focus on the twentieth century. (Contemporary visionaries frequently reference the ancient world’s lost Library of Alexandria as a historic predecessor; the desire to link the vision across the ages is also seen in the announcement of the new Bibliotheca Alexandria hosting the Wikimania 2008 conference.<sup>10</sup>) In the following pages I touch upon important moments in the pursuit of the universal encyclopedia, a technologically inspired reference work with progressive intentions, from early documentalists to Wikipedia; along the way I also ask why it took so long for this vision to become a reality.

### The Index Card and Microfilm

The idea of a personal encyclopedic device is frequently attributed to Vannevar Bush (1890–1974), an electrical engineer and advocate of America’s war research program. In a 1945 article entitled “As We May Think,” he famously outlined the idea for a *memex*, an “enlarged intimate supplement” to memory. This was envisioned as an electromechanical microfilm 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.” Noting that for a nickel, the *Britannica* could be placed on 8.5-by-11-inch microfilm and mailed anywhere for a cent, Bush predicted

“wholly new forms of encyclopedias will appear, ready-made with a mesh of associative trails running through them, ready to be dropped into the memex and there amplified.” Thus science might continue to advance without forgetting the wisdom of hard-earned experience, including science’s darker, “cruel,” applications in warfare.<sup>11</sup>

However, the memex was proposed in a larger context of lesser-known microfilm technologies and innovators. In fact, the *beginning* of the twentieth century was a seminal period for internationalism and information science. In America, Melvil Dewey (1851–1931) proffered his decimal system, founded library institutions, attempted American spelling reform, and advocated for the metric system. In Germany, Wilhelm Ostwald (1853–1932) invested much of the award from his Nobel Prize for Chemistry in *Die Brücke* (“The Bridge”), an international institute he cofounded for organizing intellectual work across the world. Emanuel Goldberg (1881–1970), once a student of Ostwald’s, advanced imaging technologies and developed a “Statistical Machine” by which microfilm records could be indexed and subsequently retrieved. In France, Suzanne Briet (1894–1989) pioneered information services at the Bibliothèque Nationale and internationally advanced the theoretical, educational, and institutional development of documentation.<sup>12</sup> Such efforts were often inspired by a new awareness of global interdependence and contemporary technologies such as the index card and microfilm. Also, early documentalists were of a time—one might even say an “age” or “movement”—when the efforts and ideas of one inspired and influenced others. One can trace Wikipedia’s heritage back to this period, as seen most clearly in the writings of two twentieth-century visionaries: Paul Otlet and H. G. Wells.

### **Paul Otlet and the Universal Bibliographic Repertory**

As a boy, the Belgian Paul Otlet (1868–1944) played at the task of extracting and organizing knowledge: he and his brother drew up a charter for a “Limited Company of Useful Knowledge.”<sup>13</sup> At the age of eighteen he wrote in his diary, “I write down everything that goes through my mind, but none of it has a sequel. At the moment there is only one thing I must do! That is, to gather together my material of all kinds, and connect in with everything else I had done up till now.”<sup>14</sup> Throughout his life Otlet was beholden to a vision of technology as a means of dissembling, synthesizing, and distributing knowledge on an international scale. Later in life, in 1918, he wrote his

vision was supported by “three great trends” of his time: “the power of associations, technological progress and the democratic orientation of institutions.”<sup>15</sup> New technologies of the day included loose-leaf binders, index cards, and microphotography. For example, Otlet and Robert Goldschmidt, an engineer and microphotography pioneer, estimated that a small can of film could hold 80,000 square meters of photographic documents such that books would soon be compact, light, permanent, inexpensive, durable, and easy to produce, conserve, and consult.<sup>16</sup> However, it was the humble 3-by-5-inch index card that would be the basis of some of his most profound insights.<sup>17</sup>

Otlet’s career began in earnest in 1891 when he joined, as a young lawyer, the Society for Social and Political Studies, under Henri La Fontaine, director of the Society’s bibliographic program and future collaborator. The span of Otlet’s career would include helping found a universal bibliographic database/encyclopedia, an international library and museum, and numerous international associations.<sup>18</sup> Furthermore, in his most famous publication of 1934, *Traité de Documentation*, he wrote of a desk in the form of a wheel from which different projects (workspaces) could be switched as they rotated—foreshadowing the multiple desktops and tabs of contemporary computer interfaces. Inspired by the arrival of radio, phonograph, cinema, and television, Otlet also posited that there were as yet many “inventions to be discovered,” including the reading and annotation of remote documents and computer speech.<sup>19</sup>

Yet there are three less speculative ideas that I think are particularly relevant to today’s Web and Wikipedia: the Repertory, the Universal Decimal Classification (UDC), and the “monographic principle.” The first two projects, of information collection and classification, arose following the first International Conference of Bibliography, in 1895, and were carried out at the newly created International Institute of Bibliography (IIB).<sup>20</sup> What began as a bibliographic “Repertory” of author files and classified subject files came to include repertories of image files (“iconographic”) in 1905 and full text files (“dossiers”) in 1907. The collections grew quickly; by 1912 the bibliographic repertory contained over nine million entries, and by 1914 the textual dossier “contained a million items in 10,000 subject files.”<sup>21</sup> Much like the millions of Wikipedia articles, and its thousands of categories and lists, we see a shared concern with collecting and ordering the world’s knowledge.

Of course, how could one possibly refer to and access all of this information? Otlet proposed a classification scheme which anticipated more recent information technologies. The UDC, based on the Dewey Decimal System, spanned over two thousand pages in its first full edition (1904–1907). Much like the URLs of today’s Web, the UDC enabled one to refer to materials deposited in the inventory of the Repertory. In addition to the relatively simple—though extensive—scheme of decimal division, Otlet complemented his system with a set of symbols specifying addition, extension, algebraic subgrouping, and language.<sup>22</sup> For example the UDC notation “069.9(100)’1851’(410.11)” specifies The Great Exhibition of 1851 in London.<sup>23</sup> This capability made the UDC more than a classification system: it was a primitive query language permitting one to specify a subset of the catalog. Indeed, a search service with documented guidelines for queries was provided until the early 1970s.<sup>24</sup> However, while being able to uniquely identify a resource on the Web today continues to be important, the way in which we manage most online information is no longer so carefully organized—and this is thought to be a feature to some, rather than a bug. As David Weinberger writes in his book *Everything Is Miscellaneous: The Power of the New Digital Disorder*, we might think of attempts at managing information in three historical orders: the order of manipulation and ordering in which physical objects are arranged (e.g., the shelving of a book); the use of metadata and the catalog (i.e., using a second-order surrogate—where we might place Otlet’s innovations); and today’s disorder (i.e., fluid, ad hoc, temporary, and disposable). Weinberger writes, “The third order takes the territory subjugated by classification and liberates it. Instead of forcing it into categories, it tags it.”<sup>25</sup> As I note in chapter 7, the notion of the “crowd” creating and managing knowledge is profoundly disturbing to some, as is a notion I think even more relevant to discourse about knowledge: the “monographic principle.”<sup>26</sup>

This primary tenet of Otlet’s schemes permitted one to “detach what the book amalgamates, to reduce all that is complex to its elements and to devote a page [or index card] to each.” Pages and cards would not be bound, but “movable, that is to say, at any moment the cards held fast by a pin or a connecting rod or any other method of conjunction can be released.”<sup>27</sup> (This is a perfect example of the ideal of “transportable knowledge” Richard Yeo speaks of with respect to the Enlightenment.) Elsewhere, Otlet wrote, “The external make-up of a book, its format, the personality

of its author are unimportant provided that its substance, its sources and its conclusions are preserved and can be made an integral part of the organization of knowledge."<sup>28</sup> Otlet envisioned being able to condense books and to strip them of their opinion so as to create "good" abstracts and even "scientific" book reviews.<sup>29</sup> As he wrote of his vision in 1903, the result, "This book, the 'Bibliion,' the Source, the permanent Encyclopedia, the Summa" would "constitute a systematic, complete current registration of all the facts relating to a particular branch of knowledge. It will be formed by linking together materials and elements scattered in all relevant publications."<sup>30</sup> This sounds much like today's Web and Wikipedia. Furthermore, Wikipedia's concern with attribution (i.e., the "Verifiability" and "No Original Research" policies) can be seen in an intention of Otlet for his own project: "Readers, abstractors, systematisers, abbreviators, summarizers and ultimately synthesizers, they will be persons whose function is not original research or the development of new knowledge or even teaching existing systematic knowledge. Rather their function will be to preserve what has been discovered, to gather in our intellectual harvests, [to] classify the elements of knowledge."<sup>31</sup>

Additionally, beyond his dedication to the technological component of the universal vision, Otlet was an internationalist and supported the foundation of the League of Nations and The International Institute for International Cooperation (which would become UNESCO) with Henri La Fontaine. Whereas La Fontaine would be recognized with a 1913 Noble Peace Prize for his international efforts, Otlet's documentation efforts were largely forgotten. Prior to World War II the Belgian government withdrew its funding, and many of the holdings were lost when the Repertory's home (at the Palais du Cinquantenaire) was occupied by German forces; over subsequent decades the collections fell into disuse and obscurity.

While the UDC is perhaps Otlet's most lasting contribution, it was not until historian Boyd Rayward "rediscovered" Otlet and argued that his vision anticipated the early hypertext of Ted Nelson's Project Xanadu that Otlet was again appreciated by those interested in the history of information science.<sup>32</sup> (And interest has become even more wide ranging: in June 2008 the *New York Times* published an article about Otlet and the establishment of a new museum and archive in Belgium.<sup>33</sup>) Rayward wrote his essay "Visions of Xanadu" in 1994, unaware of the nascent Web and that Xanadu would never be deployed. In this chapter I make a similar argument except

that I believe Otlet's Repertory foreshadowed Wikipedia. The Repertory was international, multilingual, collaborative, and predicated on technological possibility, much like Wikipedia today.

### H. G. Wells and the "World Brain"

H. G. Wells, the English novelist famous for his science fiction, was also captivated by advances in technology and the notion of a universal reference work. (While Rayward could find no evidence of direct contact between Wells and Otlet, he thinks it very likely that they at least knew of each other from their mutual attendance at the 1937 Documentation Congress in Paris.<sup>34</sup>) Like Otlet, Wells's notion of a universal reference work was not an immediate and solitary brainstorm; it was the culmination of a number of long-standing interests as prompted by the circumstances of his time. First, in his outline of a *Modern Utopia* in 1905, Wells wrote of the implications of index cards:

A little army of attendants would be at work upon this index day and night . . . constantly engaged in checking back thumb-marks and numbers, an incessant stream of information would come, of births, of deaths, of arrivals at inns, of applications to post-offices for letters, of tickets taken for long journeys, of criminal convictions, marriages, applications for public doles and the like. . . . So the inventory of the State would watch its every man and the wide world write its history as the fabric of its destiny flowed on.<sup>35</sup>

Second, since at least 1928, Wells had been advocating for an internationalist revolution, one world government, or "Open Conspiracy."<sup>36</sup> Information historian Dave Muddiman aptly identifies the key elements of this "modern" program as: "universalism and the 'World State'; planning and a central organization; a faith in scientific and technical advance, education, professionalism, expertise and benevolent socialism."<sup>37</sup> (Understandably, the idea of a "World State" that kept tabs on its citizens in such a pervasive manner can be thought to be more dystopian than otherwise and counter to our current political sentiments;<sup>38</sup> but Wells thought of these things in an optimistic light.)

Third, Wells was beginning to think of artifacts (like books) and institutions (like museums) as a type of "super-human memory" that would prompt a mental expansion for which "the only visible limit is our planet and the entire human species."<sup>39</sup> Each of these threads found their way into his 1936 proposal for a world encyclopedia, or, as he liked to call it,



a “World Brain.” (Otlet, too, at least once made reference to an “artificial brain” and Wilhelm Ostwald wrote of a “Gehirn der Welt” (World Brain) in 1912.<sup>40</sup>) Given advances in technology and the insecurity of the interwar period, Wells believed that intellectual resources were squandered: “We live in a world of unused and misapplied knowledge and skill” and “professional men of intelligence have great offerings but do not form a coherent body that can be brought to general affairs.” He hoped that a world encyclopedia could “solve the problem of that jig-saw puzzle and bring all the scattered and ineffective mental wealth of our world into something like a common understanding.” Beyond producing a resource for students and scholars, it would be an institution of “adjustment and adjudication; a clearinghouse of misunderstandings.” Ultimately, he hoped that such an institution would further the movement toward “unification and perhaps the abandonment of war.” But “without a World Encyclopedia to hold men’s minds together in a common interpretation of reality, there’s no hope whatever of anything but an accidental and transitory alleviation to any of our world troubles.”<sup>41</sup>

With respect to technology, given the resources of “micro-photography” Wells felt: “the time is close at hand when a student, in any part of the world, will be able to sit with his projector in his own study at his or her convenience to examine any book, any document, in exact replica.” And much like one of Wikipedia’s greatest strengths, it need not limit itself as a “row of volumes printed and published once and for all” but could instead be “a sort of mental clearinghouse for the mind, a depot where knowledge and ideas are received, sorted, summarized, digested, clarified, and compared” in “continual correspondence” with all that was happening in the world.<sup>42</sup> Furthermore, he proposed that the encyclopedia be in a single language (English) as it was difficult to otherwise conceive of a polyglot project satisfying his goal of social unity. Yet, it is also difficult to conceive how any such project could be genuinely universal when limited to a single language. In the case of Wikipedia, it began as an English-language work and this version remains the largest, but there are now encyclopedias in other languages. While policy for the Wikimedia projects at large continues to be discussed on the English-language email lists and the Meta wiki, the different language communities are largely autonomous. (However, multi-lingual participants do often participate across different projects, and encyclopedic articles in one language can now link to their alternative language

versions.) For Wells, perhaps English was only the initial, expedient step as he expected a universal (English-like) language would ultimately prevail in his “Modern Utopia.”<sup>43</sup>

In keeping with the universal vision, and anticipating a key Wikipedia norm, H. G. Wells was concerned that his World Brain be an “encyclopedia appealing to all mankind,” and therefore it must remain open to corrective criticism, be skeptical of myths (no matter how “venerated”) and guard against “narrowing propaganda.”<sup>44</sup> This strikes me as similar to the pluralism inherent in the Wikipedia “Neutral Point of View” goal of “representing significant views fairly, proportionately, and without bias.”<sup>45</sup> And even going beyond Wikipedia’s “No Original Research” policy, Wells thought the World Brain repository “should consist of selections, quotations, and abstracts as assembled by authorities—one need not create summaries.”<sup>46</sup> However, one must note that his propensity to write with others’ works close at hand, as he admits to with respect to the *Britannica*, perhaps led to plagiarism that was also indicative of larger character faults.<sup>47</sup>

In any case, despite claims of plagiarism, character faults, and his utopian expectations that people find alternatively progressive and frightening, Wells was a dedicated internationalist and forever looking toward the future. Like La Fontaine and Otlet, Wells thought the examples of The League of Nations, the Committee on Intellectual Cooperation in Paris, and the World Congress of Documentation were models, in spirit and application, for his own project. Yet unlike Otlet’s efforts, which were well known in their time, the World Brain never materialized beyond the ardent vision of an author.

## Digital Computers and Networks

I argue that in the first half of the twentieth century (via the examples of Otlet and Wells) we can discern a technologically inspired vision of a universal encyclopedia. This vision included collaborative capabilities—or, as Vannevar Bush spoke of, “amplifying” the contributions of others. For Otlet and Wells this collaboration was also part of their internationalist commitment. “Madame Documentation” Suzanne Briet captured this sentiment when she wrote of her library’s reading room of three hundred patrons: “peaceful with their books. Peace through books.”<sup>48</sup> Even Bush, an

architect of the atomic weapons program, hoped a better (machine-aided) memory would not let us forget the horrors of war. Yet, in the first half of the twentieth century, these visions were never satisfactorily fulfilled. For one, microfilm wasn't up to the task. As Anthony West wrote in a biography of his father H. G. Wells, "he saw before too long the technology for the storage and retrieval systems that such a thing would require was still lacking, and that its time had not yet come."<sup>49</sup> However, in the latter half of the twentieth century a new technology, the computer network, engendered new possibilities and thus inspired new directions in the creation of encyclopedias. And while the expectation that a networked encyclopedia would herald a new era of world peace lessened, the likelihood of a widely accessible and collaborative encyclopedia increased.<sup>50</sup> Even so, why did it take so long for the vision of "wholly new forms of encyclopedias" to be realized in the form of Wikipedia? In this section I use moments from the history of hypertext and digital networks to argue that it required an alignment of a coherent goal, technical practicality, and serendipity: vision, pragmatics, and happenstance.

### Project Xanadu

One ancestral line of Wikipedia's digital lineage is that of hypertext, including Ted Nelson's Project Xanadu, "the original hypertext and interactive multimedia system." Nelson's initial 1960 design "showed two screen windows connected by visible lines, pointing from parts of an object in one window to corresponding parts of an object in another window."<sup>51</sup> (Because Nelson coined the terms *hypertext* and *hypermedia*, Xanadu naturally is the first system to be labeled as such, though as Nelson himself notes, "Douglas Engelbart's NLS system at Stanford Research Institute was really the first hypertext system." Engelbart was the inventor of "word processing, outline processing, screen windows, the mouse, [and] the text link"<sup>52</sup>—and is discussed briefly in the next chapter.) Since then, Project Xanadu has had a complicated history of redesigns and attempts at commercial viability. Yet despite such difficulties, as conceived there were significant parallels between this work and its predecessors.<sup>53</sup> Its very name is a playful reference to the more fanciful aspirations of the universal vision. (The name was chosen in honor of Samuel Taylor Coleridge's unfinished poem "Kubla Khan."<sup>54</sup>) And the influence of Bush's memex on Nelson is

evidenced by Bush's essay appearing as an appendix in Nelson's seminal *Literary Machines*.<sup>55</sup> Even the index card appears in journalist Kevin Kelly's story of meeting Nelson in 1984:

Wearing a ballpoint pen on a string around his neck, he told me—way too earnestly for a bar at 4 o'clock in the afternoon— about his scheme for organizing all the knowledge of humanity. Salvation lay in cutting up 3 x 5 cards, of which he had plenty. . . . He spoke of “transclusion” and “intertwingularity” as he described the grand utopian benefits of his embedded structure. It was going to save the world from stupidity.<sup>56</sup>

This utopian characterization is not a journalist's fancy. In his book *Literary Machines* Nelson called for the support of “brilliant people looking for adventure and a challenge.” He declared: “We have to save mankind from an almost certain and immediately approaching doom through the application, expansion and dissemination of intelligence. Not artificial, but the human kind.”<sup>57</sup> This apocalyptic fear is echoed in H. G. Wells's parallel call “for a gigantic effort to pull together the mind of the race before it is altogether too late” as he feared that without “an educational revolution, a new Encyclopedism . . . we shall, as humanity, perish.”<sup>58</sup>

However, as the Web gained popular attention, Nelson and his supporters came to feel the potential and priority of the vision he had been advocating for two decades were not satisfactorily respected. Indeed, an irony of the universal vision is that its proponents are often unfamiliar with their predecessors and disappointed with their successors.<sup>59</sup> So, when the Web began its precipitous growth, largely incognizant of Xanadu, Nelson and his colleagues developed a defensive attitude toward this upstart and their own portrayal in the press. A 1995 *Wired* article entitled “The Curse of Xanadu” prompted a particularly irked and detailed response from Nelson.<sup>60</sup> This contention between proponents of Xanadu and those of the Web has quieted to some extent in recent years because of the acknowledged dominance of the Web and the honors now accorded to Nelson. Even so, Project Xanadu's mission statement still reads: “The World Wide Web (another imitation of paper) trivializes our original hypertext model with one-way ever-breaking links and no management of version or contents. WE FIGHT ON.”<sup>61</sup> In any case, while Project Xanadu inspired a generation of sophisticated desktop hypertext systems,<sup>62</sup> and the Web became the ubiquitous—if limited—networked system we use today, it was the wiki that made a dynamic and versioned Internet hypertext system widely available.

### Project Gutenberg

A second line of digital lineage originates in Project Gutenberg. Started roughly at the same time as Xanadu, Gutenberg's mission is the provision of free ebooks. Whereas Xanadu was focused on innovative ways of interfacing with information, Michael Hart, a student at the University of Illinois, started Gutenberg in 1971 to provide online access to existing print information. The story of its birth is rendered in almost mythical terms. Through friends Hart gained access to a Xerox Sigma Five main-frame computer at the university's Materials Research Lab; such a machine was extraordinarily expensive, and consequently, access to it was a valuable privilege. In fact, many of Project Gutenberg's introductory materials stress that such access was worth hundreds of thousands if not millions of dollars: "At any rate, Michael decided there was nothing he could do, in the way of 'normal computing,' that would repay the huge value of the computer time he had been given . . . so he had to create \$100,000,000 worth of value in some other manner."<sup>63</sup>

Envisioning a time when computers would be widely accessible—indeed, this computer was one of the first twenty-three that would become the Internet<sup>64</sup>—Hart began typing in a copy of the United States Declaration of Independence he happened to have in his backpack. And so "Project Gutenberg was born as Michael stated that he had 'earned' the \$100,000,000 because a copy of the Declaration of Independence would eventually be an electronic fixture in the computer libraries of 100,000,000 of the computer users of the future."<sup>65</sup> Beyond being one of the first free publicly accessible cultural resources on the Internet, Project Gutenberg is relevant to the history of the universal encyclopedic vision and Wikipedia for two additional reasons.

Initial contributions to Project Gutenberg were like Hart's inputting of the Declaration of Independence: a single contributor typing in the whole text. Whereas in one profile of the project it is claimed that Hart typed in the first one hundred books,<sup>66</sup> Hart recalls, "I had plenty of help, even back in those days, though it was mostly anonymous, and even I did not know who typed most of the first dozen or two that I didn't do."<sup>67</sup> This was laborious work, and in time the majority of texts being submitted were scanned and interpreted by optical character recognition (OCR) software. Yet this is an imperfect technology because books age and typefaces can be varied. The greatest challenge to Project Gutenberg was how to apportion and coordinate the work of volunteers who might have enough time to correct

a chapter's worth of work, but not a whole book. In 2000, Charles Frank launched Distributed Proofreaders, a complementary project to Gutenberg that would "allow several proofreaders to be working on the same book at the same time, each proofreading on different pages."<sup>68</sup> Each page of the work undergoes two proofreadings that are reconciled by a "post-processor." The importance of distributed proofreading is that it permits massive collaboration. Research on Free and Open Source Software (FOSS) development has identified this characteristic of content production as asynchronous and incremental "micro-contributions."<sup>69</sup> It is also cumulative, as Ted Nelson noted about hypertext: data can be "reorganized constantly without losing their previous organization" and therefore "order becomes cumulative—unlike most computer systems."<sup>70</sup> Indeed, Distributed Proofreaders' maxim is "a page a day"—but on average readers proof more than that. This feature of allowing many contributors to produce overlapping work in bite-sized chunks—though often becoming a consuming passion—is a powerful motif in what I call *open content communities*, those communities that openly produce software and other content, such as Wikipedia.

Project Gutenberg was also responsible for one of the first publicly available reference works on the Internet, or at least part of it: volume 1 of the 1911 *Encyclopædia Britannica* (EB11). In January 1995 Project Gutenberg published the first volume of EB11, which had passed into the public domain. However, the work then stalled. When the question of resuming was raised at the end of 2002, the resulting discussion touched on the difficulties of the work including where to get the source material, how to represent textual structure, whether to preserve illustrations, and how to deal with font difficulties. Most obviously, the project would have to accommodate the size of the text, both in the amount of material, which no one would want to type, and page size, since not all editions could fit in most scanners of the time.<sup>71</sup> (Unfortunately, efficiently scanning books often required "destructive scanning," or removing the pages from the binding, an unpalatable task to perform on historic editions. Google's scanning project uses cameras from multiple perspectives to create a 3-D model of a naturally opened book, and then "de-warps" the image of the page so it appears flat.<sup>72</sup>) Yet, the work on EB11 was resumed by Distributed Proofreaders when part 1 of volume 2 was posted with much fanfare in October 2004. Continuing with the mythic character of its origins, this event was characterized as the long-awaited return to an ancient struggle:

On the morning of October 8, 2004, near his library window overlooking a quiet lake in upstate New York, David Widger ran a series of final checks and verifications on a partitioned element of the 11th edition of the *Encyclopædia Britannica*. Yes, that same EB11 which has long been known as a formidable processing challenge throughout the Project Gutenberg community. This latest approach towards its digital conversion did little to diminish that reputation. . . . This “slice” of EB11 was not simply another single project being posted to the PG shelves, but the final component in a varied and impressive collection [that marked] the completion of Distributed Proofreaders’ 5,000th unique title produced for Project Gutenberg and the digital public domain.<sup>73</sup>

Outside of Project Gutenberg, questions of how to incorporate EB11 into Wikipedia—and even the Interpedia, a pre-Web predecessor—also proved difficult. While Gutenberg had not yet completed the task, in 2002 all twenty-nine volumes of EB11 were published at <http://1911encyclopedia.org/>. Some saw this as an opportunity to populate Wikipedia with high quality materials: EB11 was considered one of the best references of Western knowledge at the start of the twentieth century, even if rather dated by its end. Yet copyright, trademark, and substantive issues were to hinder any efforts to make use of this online version. The organization that published the twenty-nine volumes online claimed a copyright in the work it posted, arguing that its edition was an improvement upon a public domain work. Additionally, even if the text was now in the public domain the name *Encyclopædia Britannica* remained a trademark. For this reason, the Project Gutenberg version is referred to as the Gutenberg Encyclopedia. Yet even the terms of the Gutenberg Encyclopedia proved to be confusing to some Wikipedians who wished to cite the source of the work (*Britannica* or Gutenberg) without violating trademarks and their associated licenses. And substantively, some thought that any material from a 1911 work was of little use, even for historical subjects. While some material was imported as a starting point for subsequent editing, these difficulties and the extraordinary growth of home-grown content on Wikipedia rendered the issue relatively moot.<sup>74</sup>

Aside from the two obvious connections between Project Gutenberg and Wikipedia, there is a lesson here central to a theme of this chapter. A strength of Project Gutenberg was that the simple vision of sharing accessible ebooks was directly satisfied by technology available at the time: one could type existing public domain books into a networked computer using “plain vanilla ASCII.” ASCII is the legacy standard for representing

the Roman alphabet and Arabic numerals; it was the character set used by most early computer and network developers; it is still in use today. However it has no representation for accented characters, much less non-Roman scripts. Also, a file of ASCII characters is rather sedentary. No underlines, italics, or boldface—Project Gutenberg represents all of these as uppercase. Nor does ASCII accommodate links or other hypertextual innovations.

The term *plain vanilla ASCII*, is repeated in full, like a mantra, in Project Gutenberg materials. Michael Hart was well known for his opposition to any exclusive reliance upon more sophisticated textual representations such as PDF or HTML: documents, with few exceptions, must at least be available in “plain vanilla ASCII” which could then be complemented by other formats.<sup>75</sup> While frustrating to some, this insistence may have prevented the project from becoming ensnared in endless debates about formats and permitted it to achieve the success it has. As one Gutenberg participant put it:

The heart and soul of project gutenber is the plain-text file. over the years, it has been scorned and even attacked outright. some people say it's ugly. and it's far too low-tech for others. but somehow, it has survived and even thrived in a way that no other e-book technology ever has. in the process, i have grown to appreciate its tenacity, and grasp its inner beauty. this thread is for those having a love-affair with plain-text.<sup>76</sup>

However, this success had not been able to yield a complete and free online encyclopedia.

### Interpedia

Unlike Project Gutenberg, the Interpedia project was conceived of as an encyclopedia, but this conceptualization was confused by a plethora of technical options. The Interpedia Frequently Asked Questions (FAQ) document introduces the project by noting a resurgence in the early 1990s of the notion of a freely accessible encyclopedia:

According to Michael Hart the idea for a net encyclopedia has been around nearly as long as the net, at least back to 1969–71. This recent burst of activity is the result of a post to several newsgroups by Rick Gates with his idea to write a new encyclopedia, place it in the public-domain, and make it available over the Internet. Among the first responses to Rick's message was one by Gord Nickerson who suggested that this Internet Encyclopedia be fully hypertexted using a markup language such as html.<sup>77</sup>

In October 1993, when the project was proposed by Rick Gates on the alt.Internet.services Usenet newsgroup,<sup>78</sup> Internet usage was reaching a critical mass. Nontechnical members of universities and technology companies



were beginning to use email and Usenet. Computer hobbyists who typically communicated via dial-up bulletin board systems were developing Internet gateways so they too could access the Internet. And most importantly, new applications, and their network protocol and document formats, were proliferating. In addition to FTP (file transfer protocol) for transferring and storing files, email correspondence, and Usenet discussion groups, three new technologies were vying to be the next prominent Internet service. WAIS (Wide Area Information Server) retrieved documents based on keyword queries.<sup>79</sup> Gopher permitted one to browse information using menu traversal, to dig down into a publisher's taxonomy from general to specific. And, of course, there is the Web. The conundrum of which system to use is apparent in Interpedia's FAQ:

The gopher system is widely available but is not sufficiently easy to use to satisfy many people, and it does not support hypertext. Perhaps gopher software could be improved, but it doesn't seem appropriate yet. The WWW has many advantages over earlier approaches (e.g. gopher), but is not to everyone's liking. Many people do not like navigating around in hypertext, and insist that an encyclopedia must provide keyword and/or alphabetical access. Perhaps the WWW could be improved to support the Interpedia project, but it doesn't seem quite appropriate yet. It might be a good starting point though.<sup>80</sup>

It is important to remember that while it is difficult today to conceive of using the Web without a search engine, this was not an integral part of the original Web when search engines still were limited, experimental services. In addition to the confounding array of options, Doug Wilson, maintainer of the Interpedia FAQ, wrote, "the term Interpedia is ambiguous—to some it means the text, to some software, and to others what we will have when we have both." A consequence, in part, of this technical uncertainty was an ambiguity in vision. Would Interpedia be part of the Internet, or, if it references existing services, would it be something "that ends up \*being\* the net"? This confusion is further demonstrated in answer to the question on other parallel projects to Interpedia, including FAQ documents, FTP- and Gopher-based resource guides, collections of electronic art, and the Web itself.<sup>81</sup>

For about half a year Interpedia participants were relatively active on the mailing lists and Usenet group. Yet, perhaps because of these ambiguities and the explosive growth of the Web, the project never left the planning stage. Even so, it is of interest for three reasons. First, in response to the hypertextual identity crisis (i.e., are we part of the Web, or is the Web

part of us?), project participants envisioned at least a core or default set of encyclopedic articles. Articles could be submitted by anyone and quality and legitimacy would be arbitrated by a collection of decentralized seal-of-approval systems. No acceptance or rejections were necessary; instead, a seal “indicates that some article is good” and would be used by both people and the software to govern the accessibility of articles.<sup>82</sup> Second, the EB11 also proved to be a source of controversy as a strategy for initially populating Interpedia. (Michael Hart was an Interpedia member and other members eagerly anticipated all twenty-nine online volumes of the Gutenberg Encyclopedia. As noted, the first and only volume of the 1990s was posted in January 1995). Third, the process and culture of Interpedia would be facilitated by editors, whose responsibilities were “to act in good faith in the advancement of the Interpedia.”<sup>83</sup> This notion of contributors acting in good faith anticipated a cultural norm that I argue is central to Wikipedia’s collaborative culture.

### **Distributed Encyclopedia**

Though the actual Interpedia project fizzled, its newsgroup continued to be a forum for the occasional question or announcement for years to come. The notion of an Internet-based encyclopedia was no longer novel, and as the 1990s progressed the Web became the obvious platform for any such project. In hindsight, the formation of such a reference work seems inevitable. Yet, at the time, there was little clarity on how such a project would work. Ideas and half-starts came and went—or as Foster Stockwell, a historian of reference works, noted in explaining why he didn’t concern himself much with online works in 2001, they are “here today gone tomorrow.”<sup>84</sup> In 1997, Jorn Barger posted a message entitled “Beyond the Interpedia” to the newsgroup. He wrote, “from time to time, people ask if the Interpedia project—to get a full, free Encyclopedia on the net in some form—is still happening anywhere.”<sup>85</sup> The “closest descendent” known to Barger was the Distributed Encyclopedia.

Beyond this newsgroup posting, there are very few references to this project on the Web today. Its project pages themselves can only be found in the Internet Archive and do not give the impression of being more than a manifesto of a very small, if not single, number of authors.<sup>86</sup> Still, the project’s introduction clearly reflects a stabilization in how such a project was conceived. It would benefit from many contributions and it would be

distributed, meaning there would be no central authority (beyond simple stylistic conventions) or repository: each article would be hosted by the author and linked via a central index.

The irony here is that while it became clear that the Web would play a fundamental role, and an enormous strength of the Web is its hypertextual and decentralized character, Wikipedia itself is not decentralized in this way. It is not a collection of articles, each written by a single author, strewn across the Web. Instead, many authors can collaborate on a single article, stored in a central database that permits easy versioning, formatting, and stylistic presentation. Furthermore, there is a vibrant common culture among Wikipedians that contributes to Wikipedia's coherence.

### Nupedia

In January 2000, a few months prior to the first emails to the Nupedia list with which I opened this chapter, Larry Sanger emailed Jimmy Wales with a proposal. Wales, an Internet enthusiast since his days of playing in multi-user-dungeons (MUDs) in college,<sup>87</sup> had been toying with the idea of an Internet encyclopedia. When Sanger emailed him about a bloglike successor to “Sanger and Shannon’s Review of Y2K News Reports”—Y2K passed without much incident and both Sanger and Shannon were looking for new (funded/sponsored) activities<sup>88</sup>—Wales counter-proposed his encyclopedia idea and asked Sanger if he would be interested in leading the project. Each man’s career path made for a fruitful collaborative potential. Wales obtained bachelor’s and master’s degrees in finance and took courses in the Ph.D. programs at the University of Alabama and Indiana University, but never wrote a dissertation; he instead turned to the marketplace as a futures/options trader. During the explosive growth of the Internet, Wales also began investing in, and founded his own, Internet business. Sanger was a doctoral candidate in philosophy finishing his dissertation on “Epistemic Circularity.”<sup>89</sup> (This topic was to influence Sanger’s approach to addressing issues of bias and neutrality in both the Nupedia and Wikipedia.) Both men were well educated, comfortable with technology, familiar with the norms of online community and discussion, and between them had the financial, philosophical, and academic resources to launch and sustain such a project.

In February 2000, Sanger moved to San Diego to start work at Bomis, Wales’s Internet portal company. In the months before the March 9 public announcement, Sanger drafted many ideas and policies in discussion with

Wales and another Bomis partner, Tim Shell, about how to run Nupedia.<sup>90</sup> In the March 10 *PC World* article about the launch, the project was presented as ambitious and in need of contributors:

The site's managers are seeking contributors and editors with expertise in, well, almost anything. The contributors will provide the diverse content, which will be offered free of charge to both consumers and businesses. Anyone is welcome to peruse Nupedia, and any other Web site may post Nupedia's content on its own. They need only to credit Nupedia as the source.<sup>91</sup>

The article also notes that Nupedia was inspired by other open source projects like Linux and the Open Directory Project; the goal was to be open to all expert contribution and free of charge to all users, and Sanger's quoted aspiration was for Nupedia to become "the world's largest Encyclopedia." Similarly, the signature appended to the very first email sent to the Nupedia list states "Nupedia.com building the finest encyclopedia in the history of humankind."<sup>92</sup>

Unlike the Interpedia—and certainly the Distributed Encyclopedia—Nupedia shows the benefit of the resources of Wales (Bomis) and efforts of Sanger. Wales wrote to the Nupedia list: "The company behind Nupedia, Bomis, Inc., has a great deal of experience designing and promoting high-traffic websites. We intend to put that experience (and the profit from that!) behind the Nupedia project to insure that it is a success."<sup>93</sup> In the course of its first year in 2000, Sanger was the picture of frenzied cheerleading activity. In March, Sanger reported the project had 602 members and of the 140 who had filled out membership forms "about 25–40% of these (or 35–56) are Ph.D.'s or otherwise clearly bona fide experts."<sup>94</sup> By the summer the first article (on atonality) was formally published and the Advisory Board was in place. By November version 3.31 of the Nupedia.com "Editorial Policy Guidelines" was published.<sup>95</sup> Software was frequently updated throughout the year. And, throughout, Sanger was always trying to recruit new members, including the offering of T-shirts and coffee cups, and an end-of-year membership drive with cash prizes. By January 2001 there were approximately two thousand people on the Nupedia email list.<sup>96</sup>

Despite these efforts and progress, Nupedia was struggling. The recruitment efforts are evidence of the difficulty in procuring commitments from volunteers for the significant work entailed in writing an article and seeing it through the complex Nupedia editorial process.<sup>97</sup> The universal vision, this time in the form of a "dream" of a low-cost encyclopedia available to

“schoolhouses across the world” seemed reasonable, certainly compared to earlier hopes for world peace. The technology, too, seemed capable of inexpensively supplying information throughout the world, and even facilitating the work of distant contributors. Yet something more was needed and it would only be found by (seeming) accident. But before I turn to the wiki, there’s one more stop.

### **GNUPedia/GNE**

In January 2001, the same month in which the Nupedia mailing list had reached approximately two thousand subscribers, a controversy erupted around a Slashdot posting entitled “Will The Real Nupedia Please Stand Up?”<sup>98</sup> Richard Stallman, father of the Free Software movement, which itself was an inspiration for Nupedia, announced a competing project led by Hector Arena. Under the aegis of Stallman’s GNU organization the GNUPedia would implement a proposal Stallman had drafted in 1999 for a “free universal encyclopedia and learning resource.” (GNU is a recursive acronym for “GNU’s Not Unix” and it set out to replace the proprietary Unix operating system with a similar but free system.) Stallman’s proposal for a “free universal encyclopedia” had been presented in various venues in 1999 (e.g., the SIGCSE conference in March and the MacArthur Fellows Reunion in October<sup>99</sup>), but only came to be known publicly when it was made available on the Web as part of the controversial GNUPedia project announcement in 2001. Stallman outlined a vision of single-author articles distributed throughout the Web but indexed by the central project<sup>100</sup>—much like the Distributed Encyclopedia. This vision purposely eschewed any type of central authority besides a commitment to freedom, meaning any article to be linked must satisfy the criteria of permitting universal access, mirroring, modification, translation, and quotation with attribution. Given the lack of central control, these criteria would be enforced by compliant articles or indexes refusing to link to any encumbered article. Additionally, Stallman encouraged contributions from educators (whose disciplines he thought were becoming increasingly commercialized), and envisioned peer review and endorsements—similar to Interpedia seals of approval. (On the quality-ranking front, in May 2008 the German Wikipedia began using the “Flagged Revisions” feature to mark acceptable/stable versions of a page; other language editions may follow.)<sup>101</sup> In Stallman’s proposal, the Web-like assumption of decentralization was again present. And “freedom” was

ensured by the same reciprocity required by copyright licensees that govern free software: nonfree is kept separate from the world of the free. Most importantly, the proposal recognized important challenges previous projects failed to meet: contributors should appreciate that “small steps will do the job” when one “takes the long view.”<sup>102</sup>

Even so, this humble and ambitious sentiment of the tortoise getting there in the end wasn’t enough; an actual system was never realized. Because the name and the announcement were not meant to intentionally interfere with Nupedia, GNUPedia refocused as a “library of options” or “knowledgebase” and changed its name to GNE, a recursive acronym, like GNU, standing for “GNE is Not an Encyclopedia.” Stallman wrote to me that this incident was a simple case of confusion as he was in discussion with multiple people about encyclopedic projects without remembering that they were distinct, but he wanted to ensure any and all such projects would respect freedom.<sup>103</sup> Yet while GNE project participants wrestled with their new purpose, at the same time expressing concern about the centralization and complexity of the Nupedia process, Wikipedia quickly overtook both.

### The Web and Wikis

To understand the success of Wikipedia as the most credible realization of the universal encyclopedic vision, one must also understand a failing of the Web as we know it, but not as it was first conceived. In his memoir, World Wide Web inventor Tim Berners-Lee writes that his motivation was to design the Web as “a universal medium for sharing information.”<sup>104</sup> While hypertext pioneer Ted Nelson considered the Web a hobbled upstart relative to Project Xanadu, the Web as it works today falls short of even Berners-Lee’s original vision, which he now refers to in its richer potential as the “Semantic Web.”<sup>105</sup>

In any case, despite the Web’s early limitations, or perhaps because of them, in January 1993 there were nearly fifty different web browsers.<sup>106</sup> These were inspired by Berners-Lee’s original Web client and roughly implemented the specifications for HTTP (network transport), HTML (content markup), and URL (resource locators/identifiers). However, one client was to stand out among others: Mosaic, which led to Netscape. Unfortunately, some Mosaic developers were seemingly intent on overshadowing the World Wide Web and failed to implement the critical feature of

editing a Web page. Berners-Lee writes, “Marc and Eric [Mosaic developers] explained that they had looked at that option and concluded that it was just impossible. It can’t be done. This was news to me, since I had already done it with the World Wide Web [client] on the NeXT—though admittedly for a simpler version of HTML.”<sup>107</sup>

Consequently, for many people the Web became a browsing-only medium unless they were savvy enough to know how to manually publish Web pages, or fortunate enough to use a fully featured Web client such as Arena or AOLPress. Until, that is, the WikiWikiWeb. As already noted, “wiki wiki” means “super fast” in the Hawaiian language, and Ward Cunningham chose this name for his wiki project in 1995 to indicate the ease with which one could edit Web pages. Wiki makes this possible by placing a simple editor within a Web page form, with formatting and linking functions carried out by the wiki server.

At the beginning of January 2001 frustration increased over Nupedia productivity. The need to publish more articles, as well as a greater popular interest in contributing, was not well matched by the expert-dependent multistep editorial process. Hence, the stage was set for the introduction of a wiki. On January 2, at a San Diego taco stand, Sanger had dinner with Ben Kovitz, a friend from Internet philosophy lists, during which Kovitz introduced the idea of wikis to Sanger.<sup>108</sup> (The background of wiki is further discussed in the next chapter.) The wiki could be a possible remedy to Nupedia’s problems, permitting wider contribution and collaboration on articles that would then be fed to Nupedia’s editorial review. Within a day, Sanger proposed the idea to Wales and Nupedia’s wiki was announced on January 10 in a message entitled “Let’s make a Wiki”:

No, this is not an indecent proposal. It’s an idea to add a little feature to Nupedia. Jimmy Wales thinks that many people might find the idea objectionable, but I think not. . . .

As to Nupedia’s use of a wiki, this is the ULTIMATE “open” and simple format for developing content. We have occasionally bandied about ideas for simpler, more open projects to either replace or supplement Nupedia. It seems to me wikis can be implemented practically instantly, need very little maintenance, and in general are very low-risk. They’re also a potentially great source for content. So there’s little downside, as far as I can see. . . . If a wiki article got to a high level it could be put into the regular Nupedia editorial process. . . . On the front page of the Nupedia wiki we’d make it ABSOLUTELY clear that this is experimental, that Nupedia editors don’t have control of what goes on here, and that the quality of articles, discussion, etc.,

should not be taken as a reflection of the quality of articles, review, etc. on the main part of the Nupedia website.<sup>109</sup>

However, Nupedia contributors resisted Nupedia being associated with a Web site in the wiki format. Therefore, the new project was given the name “Wikipedia” and launched on its own address, Wikipedia.com, on January 15, 2001.<sup>110</sup>

## Wikipedia

Since its start, Wikipedia’s growth has been extraordinary. Within six months Sanger announced that “the Wikipedia is now useful.”<sup>111</sup> In September he proclaimed on Usenet that the “Interpedia is dead—long-live the Wikipedia”: “Interpedia’s noble dream of creating a free, open encyclopedia lives on—not quite in the form imagined, but in a ‘very’ open and free form with which many early participants would probably approve.”<sup>112</sup> Wikipedia proved to be so successful that when the server hosting Nupedia crashed in September 2003 (with little more than twenty-four complete articles and seventy-four more in progress) it was never restored.<sup>113</sup> As already mentioned, there are now scores of active language encyclopedias, millions of articles, and a handful of other Wikimedia projects. There are, of course, thousands of other wikis, many quite specialized and a few continuing forward with the universal vision. For example, the home page of the (seemingly dormant) Collective Problem Solving Wiki strikes me as true to the aspirations of H. G. Wells: “Our world has complex and urgent problems that need to be addressed. We believe there are innovative ways of solving them together online.”<sup>114</sup>

And while Wikipedia is a remarkable realization of a century-old vision, the end of this story is not as happy as it might otherwise be—nor is it really the end, just where I finish this part of the tale. In the first year of Wikipedia’s life, its radical openness and explosive growth were never reconciled with Nupedia’s goal of an authoritative expert-driven reference work. Once it was clear that a wiki could be useful, Sanger tried to introduce the idea again for Nupedia:

But by the summer of 2001, I was able to propose, get accepted (with very lukewarm support), and install something we called the Nupedia Chalkboard, a wiki which was to be closely managed by Nupedia’s staff. It was to be both a simpler way to develop encyclopedia articles for Nupedia, and a way to import articles from Wikipedia. No doubt due to lingering disdain for the wiki idea—which at the time was still very



much unproven—the Chalkboard went largely unused. The general public simply used Wikipedia if they wanted to write articles in a wiki format, while perhaps most Nupedia editors and peer reviewers were not persuaded that the Chalkboard was necessary or useful.<sup>115</sup>

Stretched between continuing frustration with Nupedia's progress, problems with unruly Wikipedians, and a widening gap between the two, Sanger failed to save Nupedia and alienated some Wikipedians who saw his actions as increasingly autocratic.<sup>116</sup> Additionally, with the burst of the Internet bubble, Sanger, among many others in the industry, was laid off from Bomis and resigned from his Wikipedia role shortly thereafter. Sanger's subsequent commentary from the sidelines, particularly his continued criticism of Wikipedia not respecting the authority of experts, prompted additional negativity toward him. In April 2005, Sanger published his memoirs of Nupedia and Wikipedia, which sparked a controversy over whether Sanger even deserved credit as a cofounder of Wikipedia.<sup>117</sup> In March 2007 Sanger launched a new encyclopedic project, Citizendium, with the intention of improving "on the Wikipedia model by providing 'reliable' and high-quality content; . . . by requiring all contributors to use their real names, by strictly moderating the project for unprofessional behaviors, and by providing what it calls 'gentle expert oversight' of everyday contributors."<sup>118</sup> In early 2009, in a sad irony for a project based on good faith, the question of credit and cofounding erupted again, with Sanger and Wales becoming even more embittered and accusing each other of dishonesty.<sup>119</sup> Sanger's exit from Wikipedia will be further touched on when I consider leadership in such communities, as will the larger social debate about "experts versus amateurs."

### **Conclusion: Predicting the Future, Reading the Past**

A history professor of mine once wisely noted "historians stink at predicting the future." Predictions about technology, regardless of who makes them, seem especially problematic.<sup>120</sup> Even those who help "make" the future are no better at prediction. In this chapter I considered those looking back, those looking forward, and those struggling in their present to implement a universal encyclopedic vision. For a long time, no one got it quite right. But people, being people, try, and try again. And that story is revealing in at least two ways.

First, even unfulfilled visions, failed projects, and erroneous predictions tell us something about those people and their time. The history recounted in this chapter speaks to the alluring and enduring notion of an ambitious project of human knowledge production and dissemination: a universal encyclopedia. This vision persisted throughout the twentieth century even though each instance was prompted by different technologies and entailed differing levels of accessibility in production: Otlet's documentalists, Wells's technocrats, Nupedia's scholars, and Wikipedia's "anyone."

Second, a question throughout this chapter is why did it take so long for the vision to be realized? A possible answer can be detected in the overlapping spheres of vision, pragmatics, and happenstance; interesting things happen when those stars align. Perhaps the best example of this can be seen in the expectation (i.e., of the Distributed Encyclopedia and GNUPEdia) that once it was clear the Web would be a platform for such an encyclopedia, it would also be decentralized. But, Wikipedia is centralized, in part, because wikis made editing the Web possible again for many people—and the loss of editing capabilities from Berners-Lee's original vision was seemingly another chance event. Wikis have other features that make them useful for an encyclopedia (e.g., versioning and simple inter-wiki linking)—though, seemingly, Wales himself thought such a notion would not be received well and Ward Cunningham predicted that the result would be more a wiki than an encyclopedia<sup>121</sup>

In any case, the projects discussed in this chapter are attempts at realizing a universal vision, encompassing the goodwill of collaborators and reaching toward global accord. While it is a mistake to argue all reference works are necessarily progressive, as I warn in chapter 7, even *Britannica*—often thought to be the conservative opponent of the *Encyclopédie* in the 1800s and Wikipedia today—shared this sentiment of global accord in a preface to a 1956 edition of its world atlas: "To the men, women, and children of the world who, by increasing their knowledge of the earth and its people, seek to understand each other's problems and through this understanding strive for a community of nations living in peace, the *Encyclopædia Britannica* dedicates this volume."<sup>122</sup> And while few would argue that Wikipedia will necessarily further world peace, in the next chapter I argue "good faith" culture is necessary to its production and an occasional consequence of participation.

### A Timeline of Events

1895 Otlet's Permanent Encyclopedia: liberating ideas from the binding of books.

1936 Wells's World Brain: a vision of a worldwide encyclopedia using microfilm.

1945 Bush's memex: a vision of a hypertextual knowledge space and new forms of encyclopedias.

1965 Nelson's Xanadu: a vision of hypertext.

1971 Hart's Project Gutenberg: a vision of providing ebooks through achievable means ("plain vanilla ASCII").

1980s Academic American Encyclopedia is made available in an online experiment; multimedia CD-ROMs soon follow.

1991 Berners-Lee's World Wide Web: a vision of highly accessible read/write.

1993 Interpedia: an ambiguous vision lost among too many infrastructural options.

1995 Cunningham's WikiWikiWeb: making the Web easy to edit collaboratively.

1999 Distributed Encyclopedia: many people should contribute independent essays that could be centrally indexed.

1999 Stallman's "The Free Universal Encyclopedia and Learning Resource."

2000 Distributed Proofreaders: distributing the task of proofreading among many.

2000 (March 9) Nupedia launched: a FOSS-inspired expert-driven free encyclopedia.

2001 (January 10) "Let's make a wiki."

2001 (January 15) [www.wikipedia.com](http://www.wikipedia.com) launched.

2001 (January 16) GNE project announced.

2001 (September) "Interpedia is dead—long live the Wikipedia."



This is a section of [doi:10.7551/mitpress/8051.001.0001](https://doi.org/10.7551/mitpress/8051.001.0001)

# Good Faith Collaboration

## The Culture of Wikipedia

By: Joseph Reagle

### Citation:

*Good Faith Collaboration: The Culture of Wikipedia*

By: Joseph Reagle

DOI: 10.7551/mitpress/8051.001.0001

ISBN (electronic): 9780262289719

Publisher: The MIT Press

Published: 2012



The MIT Press

© 2010 Joseph Michael Reagle Jr.

All rights reserved. No part of this book may be reproduced in any form by any electronic or mechanical means (including photocopying, recording, or information storage and retrieval) without permission in writing from the publisher.

For information about special quantity discounts, please email [special\\_sales@mit-press.mit.edu](mailto:special_sales@mit-press.mit.edu)

This book was set in Stone Sans and Stone Serif by the MIT Press. Printed and bound in the United States of America.

Library of Congress Cataloging-in-Publication Data

Reagle, Joseph Michael.

Good faith collaboration : the culture of Wikipedia / Joseph Michael Reagle Jr. ; foreword by Lawrence Lessig.

p. cm. — (History and foundations of information science)

Includes bibliographical references and index.

ISBN 978-0-262-01447-2 (hardcover : alk. paper) 1. Wikipedia. 2. Electronic encyclopedias—Case studies. 3. Wikis (Computer science)—Case studies. 4. Communication in learning and scholarship—Technological innovations—Case studies. 5. Authorship—Collaboration—Case studies. 6. Online social networks—Case studies. I. Title.

AE100.R43 2010

030—dc22

2009052779

10 9 8 7 6 5 4 3 2 1