

THIS ESSAY IS A CHAPTER IN THE FOLLOWING DISSERTATION: Hill, Benjamin Mako. “*Essays on Volunteer Mobilization in Peer Production*.” Ph.D. Dissertation, Massachusetts Institute of Technology, 2013.

PLEASE CHECK [THIS PAGE](#) FOR A NEWER VERSION OF THE ESSAY BEFORE CITING OR DISTRIBUTING IT.

## *Almost Wikipedia: Eight Early Encyclopedia Projects and the Mechanisms of Collective Action*

Benjamin Mako Hill ([mako@mit.edu](mailto:mako@mit.edu))

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

**Abstract:** Before Wikipedia was created in January 2001, there were seven attempts to create English-language online collaborative encyclopedia projects. Several of these attempts built sustainable communities of volunteer contributors but none achieved anything near Wikipedia’s success. Why did Wikipedia, superficially similar and a relatively late entrant, attract a community of millions and build the largest and most comprehensive compendium of human knowledge in history? Using data from interviews of these Wikipedia-like projects’ initiators and extensive archival data, I suggest three propositions for why Wikipedia succeeded in mobilizing volunteers where these other projects failed. I also present disconfirming evidence for two important alternative explanations. Synthesizing these results, I suggest that Wikipedia succeeded because its stated goal hewed closely to a widely shared concept of “encyclopedia” familiar to many potential contributors, while innovating around the process and the social organization of production.

### INTRODUCTION

Over the last decade, several books (e.g., Reagle, 2008; Tapscott and Williams, 2008; Lih, 2009) and more than 6,200 peer-reviewed articles have been published about Wikipedia.<sup>1</sup> With interest driven by the rapid growth of both its encyclopedic product and community of contributors that builds it (see Ortega, 2009), a range of scholars and journalists have held Wikipedia up as a model of collaboration and collective action on the Internet (e.g., Benkler, 2006; Shirky, 2010). Although Wikipedia was not the first website to use a wiki – a piece of server software, invented by Ward Cunningham, that allows users to create and edit Web page content using any Web browser (Leuf and Cunningham, 2001) – its example has inspired countless follow-on efforts. By 2010, the wiki-

---

<sup>1</sup>This number reflects the number of results from a *Google Scholar* search for articles with “Wikipedia” in the title. (Accessed June 16, 2010)

hosting firm Wikia alone hosted more than two thousand websites with the phrase “pedia” in their URLs – three percent of their hosted projects.<sup>2</sup>

It would be an enormous understatement to state that not all attempts at collaboration and collective action online have enjoyed Wikipedia’s success. Indeed, research has shown that few free/libre open source software projects (FLOSS) have more than one contributor (Healy and Schussman, 2003); the vast majority of wikis are similarly uncollaborative in their production (Reich et al., 2012); most remixable media never elicits a derivative (Hill and Monroy-Hernández, 2013). As a result, an important research question for scholars of peer production is: *Why do some peer production projects successfully attract contributors while most do not?* It is also one largely unaddressed in the literature on peer production to date.

Through extensive research on Wikipedia, scholars know quite a lot about *how* Wikipedia built a large community. As a group, we have documented Wikipedia’s history, organization, processes, and routines in detail. However, because we have not compared Wikipedia, for the most part, to similar failures, we still know very little about *why* Wikipedia succeeded. This paper examines seven volunteer-driven online collaborative encyclopedia projects (OCEP) founded before January 2001, when Wikipedia was launched. Although most of these projects attracted some participants, and while some built sizable groups of volunteers and substantial collections of articles, none of these projects became examples of mass collaboration on anything near the scale of Wikipedia. Using qualitative analysis of interviews from initiators of these projects and extensive archival material, I use multiple case study methods (Eisenhardt, 1989) to build support for a series of propositions that both offer a set of potential answers to the question posed by Wikipedia’s surprising success and that suggest important mechanisms of collective action in general.

Analysis of the data from initiators and archival data suggests that Wikipedia attracted a large community of contributors, while similar projects struggled, for three reasons (presented as propositions): First, (P1) Wikipedia attracted volunteer contributors because it took as its goal a model of “encyclopedia” familiar to potential contributors. Second, (P2) Wikipedia attracted volunteer contributors because it offered low barriers to contribution. Third and finally, (P3) Wikipedia attracted

---

<sup>2</sup>Results of personal research using a dataset of all of Wikia’s published database “dumps” gathered in April 2010.

		Innovativeness of Goal/Product	
		Familiar	Novel
Innovativeness of Process/Tools	Familiar	Traditional products using traditional methods and tools. <i>"Like Encyclopedia Britannica — just online and free."</i>	New products using traditional methods and tools. <i>"A new type of encyclopedia, but produced like the old ones."</i>
	Novel	Traditional products using new methods and tools. <i>"Like Encyclopedia Britannica, but produced in a radically new way."</i>	New products using novel methods and tools. <i>"A new type of encyclopedia produced in a radically new way."</i>

Figure 1: Representation of a theoretical design space in which peer production projects vary between high and low levels of innovation in their goals and products (columns) and methods and processes (rows). I propose that projects in the bottom-left shaded quadrant will be most effective.

volunteer contributors because it deemphasized attribution and social ownership of content. Before concluding, I also consider and reject two alternative explanations: first (AE1), that Wikipedia was more technologically sophisticated than alternatives; and second (AE2) that early OCEPs were simply ahead of their time. Taken together, my results suggest a theory of volunteer mobilization in peer production and beyond. Although Wikipedia was extremely innovative in terms of *how* the project set out to build its encyclopedia, it was profoundly uninnovative when it came to *what* it set out to build.

Early OCEPs varied in the familiarity of both their products and their processes. Figure 1 shows the four theoretically possible combinations of familiar (or novel) processes and products as well as providing a concise description of how an OCEP of each combination might be described. Wikipedia was the only project in my sample of OCEPs coded as falling into the bottom-left quadrant that reflects the combination of a familiar goal (e.g., “simply reproduce Encyclopedia Britannica”) with innovative methods (e.g., “anybody can edit anything”). My analysis suggests Wikipedia attracted an enormous community while very similar projects did not because Wikipedia sought to build a product that potential contributors were already deeply familiar with, while also adopting a novel set of processes, tools, and methods for organizing production. Other projects struggled both

to get contributors onto the same page and to organize production in a way that took advantage of new communication technologies.

## BACKGROUND

The last decade has seen a large and growing scholarly interest in online, volunteer-driven, collaborative projects. Although early interest focused on software, theories of FLOSS were extended to a variety of online communities and products under the theoretical umbrella of “peer production” (Benkler, 2002, 2006). The peer production literature has been heavily influenced by the author and software developer Eric Raymond (1999) who influentially argued that FLOSS is driven by contributions from a large number of self-motivated participants. According to Raymond, this collaborative process is inherently superior to proprietary development processes and leads to higher quality products.

Early academic work on FLOSS relied heavily on Raymond’s description to justify and frame FLOSS research. Citing Raymond’s descriptions of ad-hoc self-organization and large-scale voluntary contributions, Benkler (2002, 2006) offers a theory of peer production, using transaction-cost economics from Coase (1937) and Williamson (1981). Benkler suggests that peer production is made possible by decreases in communication costs caused by the Internet and suggests that the phenomena is unique from previously theorized organizational forms (i.e., markets and firms) by both its ability to aggregate many small contributors from individuals with diverse motivations and by its lack of traditional hierarchical and legal systems of control (Benkler, 2013). “Coase’s Penguin,” Benkler’s article coining the term peer production, foreshadows future work in that it used Wikipedia (then only one year old and largely unheard of) as an example of how theory from FLOSS could be used to explain the collaborative creation of a variety of information goods. His article also acts as a template for subsequent peer production research in its reliance on a small number of the most successful examples of peer production.

In their extensive literature review, Crowston et al. (2010) suggest that a large proportion of FLOSS research has been framed as case studies of a handful of hugely successful projects like Apache, Mozilla, and the Linux kernel (e.g. Krishnamurthy, 2005; Lakhani and von Hippel, 2003;

Mockus et al., 2002; MacCormack et al., 2006, 2012). Nowhere has this been more true than in regards to Wikipedia. Wikipedia is both the most visible example of peer production and the most successful example of volunteer-based productive mass-mobilization. As the fifth most-visited website in the world, the English version of Wikipedia has over 13 million registered contributors and more than 3.4 million articles.<sup>3</sup> As the subject of more than 6,200 academic articles, it is also likely the most intensely studied.

Raymond's description of FLOSS has attracted scholarly criticism (e.g., Berzroukov, 1998; Krishnamurthy, 2005). Both members of the FLOSS community (e.g., Hill, 2005) and sociologists (e.g., Lin, 2005) have tried to re-frame Raymond's work as a Utopian ideal rather than a representation of empirical reality. Large cross-sectional analyses of FLOSS projects by Krishnamurthy (2002), Healy and Schussman (2003), and Schweik and English (2012) have presented empirical evidence that calls into the question claims of FLOSS collaborativeness and has shown that participation in FLOSS projects is extremely skewed: the median number contributor to a project is one (Healy and Schussman, 2003). Research on wiki communities (Reich et al., 2012; Kittur and Kraut, 2010) and remixing communities (Hill and Monroy-Hernández, 2013; Luther et al., 2010) has shown similar struggles with mobilization in the large majority of attempts at peer production.

Some researchers have adopted alternative measures of "success" that rely on downloads or the attainment of project goals (e.g., Crowston et al., 2003; English and Schweik, 2007; Schweik and English, 2012; Luther et al., 2010). Using these metrics, most successful projects are the work of a single participant or a very small group. For example, in their comprehensive study of SourceForge projects, Schweik and English (2012) show projects that have made a series of releases and created sustained commons have an average of 1.2 contributors. Among collaborative FLOSS projects, West and O'Mahony (2005) finds that many are firm-based efforts with little or no outside contributions.

Although this research shows that uncollaborativeness need not be equated with failure, uncollaborative attempts to create online commons are failures of peer production; after all, inspired by Raymond, much of the interest in Wikipedia and FLOSS stems from the benefits of collaboration;

---

<sup>3</sup>Statistics are taken from <https://secure.wikimedia.org/wikipedia/en/wiki/Special:Statistics> (Accessed December 8, 2010).

von Hippel and von Krogh's (2003) "private collective model" assumes collectivity; advantages of Murray and O'Mahony's (2007) "cumulative innovation" are irrelevant without actors reusing and recombining ideas. Peer production requires pluralistic peers to produce cooperatively. In this sense, collaboration's rarity is a critical problem for research driven by an interest in quality that, theory suggests, is grounded in collaboration.

As a result, the fundamental research question facing FLOSS becomes one largely unaddressed in the literature to date: *Why do some peer production projects successfully attract contributors while most do not?* Answering this question involves several departures from most previous research. First, research should treat the peer production *project* as the unit of analysis. Second, research should take mobilization as its dependent variable. Third, because it may be possible to mobilize individuals without eliciting contributions, research should consider the degree to which projects are able to direct the energy of participants into useful contributions toward a project's goal.

Studies of mobilization in peer production can speak to issues and scholarship on mobilization more generally. Historically, questions of mobilization have been the domain of the sociological literature on social movements. McCarthy and Zald's (1977) resource mobilization theory (RMT) uses an analogy to firms to describe movements as constituted by social movement organizations (SMOs) competing for scarce resources – usually volunteer participants or donations – within a social movement industry (SMI). That said, although RMT has played a major role in the literature on social movements, very little empirical work within RMT has considered interactions at the SMI level. Important exceptions are the work of Minkoff (e.g., 1999) and Soule and King (2008). Additionally, it has been difficult, historically, for social movement researchers to study nascent SMOs because early-stage SMOs are not visible before they are active. In her work, Minkoff relies on organizational registers that include only established, usually incorporated, SMOs. Soule and King use a dataset of New York Times reports of protests. Empirical work in the sociological literature on mobilization has never been able to capture very early failed attempts.

Because they occur entirely online, many peer production projects leave behind "digital traces" (Wesler et al., 2008) in the form of publicly available archival data. These data are of scope and scale that was unavailable to previous scholars of mobilization. This more complete record allows re-

searchers to explore nascent attempts at movements that did not succeed in attracting *any* additional participants or that attracted participants but did not accomplish their initial goals. The following analysis considers projects in this nascent stage. Using interviews and extensive digital trace data, I build a general, if partial, theory of early stage mobilization and suggest propositions for why some attempts at collective action succeed, while most do not.

## FAILED “WIKIPEDIAS”

To construct a comprehensive sample of early OCEPs active before Wikipedia, I used characteristics cited as key qualities of the initial Wikipedia project by Reagle (2008) and others. In particular, I sought out encyclopedia projects that were English-language, collaboratively-produced, online, freely-available, and volunteer driven. A sample of projects was initially formed from consulting an existing list of online encyclopedia projects hosted on Wikipedia<sup>4</sup> and through searches of news articles on the database *Lexis Nexis*. Additionally, each interviewee was asked to list any other online encyclopedia projects that they were aware of. I am confident that my current sample represents the full population of publicly announced OCEPs from the period before Wikipedia was launched.

Before January 2001, when Wikipedia was founded, there existed seven other publicly announced attempts to create English-language volunteer-driven, online collaborative encyclopedia projects. Although not all of these projects self-identified publicly as encyclopedias, each of them sought to build general purpose encyclopedias in the sense that their proposed products can be accurately described using the Oxford English Dictionary’s definition of encyclopedia: “a literary work containing extensive information on all branches of knowledge.” Moreover, each of these projects either publicly described themselves as an encyclopedia or was identified as one in the press. During interviews, initiators of these projects confirmed that their project could accurately be described as encyclopedias. Each of these projects elicited content exclusively from volunteers on the Internet. Although some made special attempts to recruit content providers, none paid for content.<sup>5</sup> In this sense, each of the projects represented a form of crowd-sourcing or user generated content (Howe,

<sup>4</sup>[https://secure.wikimedia.org/wikipedia/en/wiki/List\\_of\\_online\\_encyclopedias](https://secure.wikimedia.org/wikipedia/en/wiki/List_of_online_encyclopedias)

<sup>5</sup>One project, *The Vines*, paid authors a portion of advertising revenue when their articles were viewed. Because they introduced paid labor which, theory suggests, may complicate questions of mobilization, I have not included them in the analysis.

Project	Total Participants	Total Articles
Interpedia	400	<50 (?)
TDEP	1 (?)	5
Everything2	50,000+ (?)	500,000+
h2g2	5,000+	13,000+
TheInfo	20 (?)	50 (?)
Nupedia	2,000+ (?)	24
GNE	300+	3-4 “test” articles
Wikipedia	500,000+	2,000,000+

Table 1: List of OCEPs started in and before January 2001. Details of the size of the projects in total contributors and total articles are shown. These include either the total size over the life of the project or, for the projects that continue today, the total number in December 2010.

2008). Each project was published on the Internet and its content was available to the public at no cost.

Several projects operated under multiple names, as indicated below. All projects are referred to using their final names throughout this article, although quotes from interviewees or archival sources have not been altered. Each OCEP in my sample is described below in the order that they were launched.

**Interpedia** (also referred to as the **Internet Encyclopedia Project**) was created in late 1993 by library school student Rick Gates. The project was passed off within its first few months to a group led by Canadian technologists Douglas Pardoe Wilson and Robert Neville. The Interpedia is the only project in my sample that effectively pre-dates the World Wide Web. Interpedia was organized over email lists and USENET where its discussion group sent over 700 messages from several hundred participants. Volunteer participants were engaged in four distinct technological projects to build clients and servers. Contributions were collected for use with these clients but were never published.

**The Distributed Encyclopedia Project (TDEP)** was founded by a German computer consultant in 1997 and was re-launched in 1998 with additional content. The project did not use a dynamic web application but was designed either to host articles emailed to a group of maintainers or to link to articles on other websites from a central server. In this sense, TDEP aimed to act as an index to encyclopedia articles across the web and even to multiple articles on the same topic. TDEP never



attracted any articles from users other than its initiator.

**h2g2** was a project funded by *The Digital Village*, a British media and video game company connected closely to the science fiction author Douglas Adams. The goal of the site was, inspired by Adams' best-selling book series, to create *The Hitchhiker's Guide To the Galaxy: Earth Edition* – an irreverent encyclopedia about earth in the style of the fictional guide. The project was initially launched live on the BBC and was later acquired by that organization. It was spun out from the BBC in 2012 and continues to operate separately today.

**The Info Network (TheInfo)** was a project of the 14 years old “wunderkind” and technology entrepreneur Aaron Swartz. Before launching TheInfo, Swartz co-authored the specification for RSS – a core technology for data syndication on the web used by nearly all dynamic websites today. TheInfo attracted news coverage and an award from online entrepreneur Philip Greenspun's organization *Ars Digita* but struggled to attract contributors. It continued with a very small following for more than four years before being taken offline by a hard disk crash.

**Nupedia** sought to be a traditional encyclopedia project available at no cost on the web. Nupedia recruited thousands of expert participants with academic credentials and subjected articles to a rigorous review process. The project was started by Jimmy Wales and Larry Sanger – the two initiators of Wikipedia – about a year before Wikipedia was launched. Wikipedia was initially started as a side project of Nupedia and Nupedia continued after Wikipedia was launched. As Wikipedia became more successful, Nupedia was increasingly ignored and eventually abandoned.

**Everything2 (E2)** was founded in 1998 as a hypertext encyclopedia project created by Internet entrepreneur Nathan Oostendorp and funded largely by the sale of the very popular technology news site Slashdot during an early period of the dotcom boom. The project was originally founded as *Everything* but was changed to E2 to reflect a major technological upgrade to the software that ran the site. E2 aimed to include more than just encyclopedic content and hosted writing in a variety of forms. The community stabilized in size and continues today with a contributor-base that is about around 5% that of Wikipedia.

GNE was started as **Gnupedia** and renamed to the recursive acronym **GNE's Not an Ency-**

**clopedia** – both references to the technology project GNU (Stallman, 2002). GNE was a project initiated by FLOSS developers inspired by an essay by Richard Stallman calling for the creation of a free encyclopedia (Stallman, 1999). The project changed its name to avoid confusion with Nupe-  
dia. Gnupedia was announced almost simultaneously with Wikipedia and initially attracted similar visibility and attention. Gnupedia continued for about a year, attracted several hundred participants who engaged in a large amount of policy and technical development, collected essentially no encyclopedic content, and eventually withered.

## DATA AND METHODS

Data used in this analysis came primarily from interviews with project initiators<sup>6</sup> and was supplemented with extensive archival data. All interviews were opened-ended and semi-structured. Interview protocols were designed to give interviewees space to reflect on their experience with their OCEPs. Toward this end, the protocol included broad questions about initiators' backgrounds and a series of questions about their OCEP's ideation, planning, audience, contributors, structure, and experience. At the end of each interview, I asked each initiator a series of questions about their impressions of Wikipedia and asked them to compare Wikipedia to their project. The full protocol is included as Appendix A.

With one exception, for which archival data is available, I was able to contact and interview at least one representative from each early OCEP. In two cases, in-person interviews were arranged. In each of the other cases, interviews were conducted over the phone or VoIP. Interviews lasted between 30 minutes and 3 hours with the median length near 90 minutes. Multiple initiators were interviewed when this was relevant and possible. All interviews – 11 in total – were recorded and fully transcribed.

Project initiators proved to be an enormously valuable source of data given my research question. In every case, informants were deeply familiar with Wikipedia, had closely followed its development, and had carefully compared their project and their own choices as initiators to Wikipedia's

---

<sup>6</sup>Because the term “founder” is a contested term for several of the projects in my dataset, I use the term “initiator” to refer to any individuals involved in conceptualizing and launching an OCEP. Most, but not all, of my informants self-identify as founders.

choices and experiences. Every single informant indicated that they considered Wikipedia a missed opportunity for themselves and their projects. As a result, each initiator had reflected, thoughtfully and at length, over a span of more than a decade, on the core research question at the heart of this study.

Additionally, I assembled extensive archival data for each project. I downloaded full email or USENET archives for Nupedia, GNE, and Interpedia which document both the planning stages and the period after “launch.” Similar material was not created or was not available for other projects. Using archival information and discussion forums available on the web and taken from the Internet Archive’s *Way Back Machine*, I gathered documentation of process discussions for E2, TheInfo, and h2g2. Finally, I did full text keywords searches for each project in Lexis Nexus and assembled lists of published news articles for each project from before 2001. The total archival dataset was more than 3,000 pages of text.

Data analysis was structured as a comparative multiple case study (see Eisenhardt, 1989). In most cases, I was able to confirm both descriptions and ascriptions by project initiators using available archival data with statements by the initiators of other projects, and, where multiple initiators were interviewed, by statements of co-initiators. All interviews were hand coded through an iterative step-wise process using the RQDA qualitative data analysis software (Huang, 2012). Using methods from grounded theory (Charmaz, 2006), codes were initially inductive and emergent from the text. Subsequently, data was also coded using deductive codes representing concepts from theory (see Corbin and Strauss, 2007). As new codes were created or merged, data was reviewed and recoded in an iterative, step-wise manner. Codes were categorized and groups were combined into higher-level meta-codes to create themes.

## RESULTS

The process of coding and code aggregation led to three distinct themes associated with propositions for why Wikipedia succeeded. Each of these propositions was raised by initiators from at least projects and these propositions reflect the full list of themes cited by at least three projects as explanations for Wikipedia’s success. These propositions are:

**P1:** Wikipedia attracted contributors because it was built around a *familiar product*.

**P2:** Wikipedia attracted contributors because it offered *low barriers to contribution*.

**P3:** Wikipedia attracted contributors because it offered low attribution and *low social ownership of content*.

Each of these propositions is discussed in depth in the subsections below.

*Proposition 1: Familiar Product*

Many respondents suggested that Wikipedia was effective in attracting contributors because it took steps to clearly define itself as an encyclopedia. On the other hand, there was considerable variation in how closely other OCEPs hewed to traditional conceptions of “encyclopedia-ness” and many projects attempted to extend the traditional encyclopedia genre to explore new possibilities for encyclopedic reference works online. Wikipedia has called itself an encyclopedia on every page since its establishment. Through this type of appeal to tradition, Wikipedia was among the most conservative projects in my sample in terms of its goal. Despite the new forms of reference works made possible by the Internet, Wikipedia set out to be just an encyclopedia.

In his book and in several articles, media theorist Joseph Reagle places Wikipedia into historical context by describing it as the latest in a long series of attempts to codify and transfer knowledge through reference works that he traces back to Diderot and the Enlightenment (Reagle, 2009, 2008; Loveland and Reagle, 2013). Reagle suggests that Wikipedia’s core policies – e.g., neutral voice, coverage of only notable subjects, a prohibition on original research, authority built through references to external published sources – can all be traced to an understanding, familiar to potential contributors, of what an encyclopedia should be and has been.

Many other OCEPs were more ambitious in the way they defined their projects’ goals. By offering subtly different ideas of *what* they were building, other OCEPs complicated and elaborated on the idea of an encyclopedia in describing their work. For example, in one of the very first articles shared on his site, E2 initiator Nate Oostendorp explained:

Everything is a flexible web database, created by [Block Stackers Inc.] which seeks to find the best way to store and link ideas. The result: it's absolutely crazy... (Oostendorp, 1999)

E2 struggled with explaining to users what it sought to be without constraining itself to existing genres. That said, Oostendorp's answer may not have been of much guidance to prospective E2 contributors trying to decide what to contribute. An early E2 "frequently asked questions" (FAQ) document included, as its very first entry, the question, "What is Everything?" In answering the question, E2's only paid editor, struggled with the difficulty of explaining to new users exactly what E2 was and was not:

Everything is what you make of it, that's the bottom line. It's open-ended, open-minded and waiting for you. You can node your diary, CD collection, dreamlog, notes on the apocalypse or a recipe for fettuccine alfredo. You can sit around and read what other people have written. You can recommend changes in the system. You can do almost anything you want to [providing] you spell it right...

Everything is an ever-growing, pulsating database that moves through cyberspace like a death-borg...slowly collecting and assimilating information and nonsense until...until... (Postma, 2000)

In interviews, multiple early E2 contributors confirmed that despite the fact that a large proportion of early content in E2 was encyclopedic in nature, users had trouble learning *what* to contribute. As a result, while the site was able to build a following of thousands of contributors, it failed to take off in the way that Wikipedia did.

In an interview, Aaron Swartz, initiator of TheInfo, recalled that his project similarly resisted a strict encyclopedic frame:

I don't think I conceived of it as like "let's just put an encyclopedia online." I think I probably [thought] like, "this is going to be an exploration and we're going to figure out what a reference work online looks like" (Interview).

Elsewhere, Swartz explained that he was highly influenced by writers like Vannever Bush and Theodore Nelson who each described a science fiction-like future of interlinked databases that went beyond what was possible in traditional encyclopedias (see Bush, 1945; Nelson, 1981). Swartz, like other initiators, saw his project as an opportunity to explore these possibilities and sought to do so by not limiting his project to existing paper-based conceptions of what a reference work should be.

Although h2g2 was more reserved in its ambitions, it reflects yet another example of an OCEP that departed from the traditional encyclopedic frame. In Douglas Adams' books, the *Hitchhikers Guide* is a electronic encyclopedia that forms a compendium of essential knowledge about the universe and that is described as being humorous and irreverent in tone – especially when compared to stodgier *Encyclopedia Galactica* which is described by Adams as a traditional encyclopedia. Avoiding the term encyclopedia, h2g2 explained at its launch that it was, “a repository of human experience and knowledge” (Archival Data). As was typical of other OCEPs, h2g2's minor deviation from the well-understood encyclopedia model caused friction with contributors who struggled to understand what it was they were being asked to contribute. One h2g2 initiator describes this difficulty in a way that was echoed by other interviewees:

So one of the problems was firstly that people would be writing completely fictional stuff about the universe. Y'know, about the Hitchhiker's universe. And we'd go, "No, no, no, no. You're not getting it. This is for real people. This is about the real world."

Then what they did, at the same time, was that they'd also do stuff about the real world, but try and write it from the point of view for an intergalactic guide. So we'd get articles about soccer that would start with, "On the planet earth which is the third planet out from the solar system Sol, the humans like to play, blah blah blah blah." Shut up. Alright? It's like, "this is going to be read by humans who live on earth."

We had piles and piles of that shit. And we had to shovel our way out from under it (Interview).

Even though h2g2 was encouraging the production of encyclopedia articles, the fact that they described their project as the *Earth Edition of The Hitchhiker's Guide to the Galaxy* confused potential

contributors. Although most potential contributors to OCEPs arrived with some familiarity with encyclopedias, every h2g2 contributor first had to learn how h2g2 was different. The quotation above, and many others like it, suggest that even when OCEPs deviated only slightly from a traditional encyclopedia in their stated goal, this deviance significantly hindered mobilization.

Like many other interviewees, E2 initiator Nate Oostendorp suggested that E2's attempt to avoid a pure encyclopedic framing was the major reason that E2 failed to mobilize contributors to the degree that Wikipedia did:

I don't think we ever used the term encyclopedia and that probably would have been smart...

Wikipedia had a much more focused purpose than Everything2. Everything2 was, just by its nature, sort of zen koan like, everyone who was involved with it thought it completely defined description. And that, I think, was to its, ultimately to its detriment.

Versus Wikipedia which was like, "we're going to be like the encyclopedia, like the World Book Encyclopedia but huge and comprehensive. But we're going to keep this impartial tone and everything has to be referenced and that sort of thing" (Interview).

We can think of this contrast between models in terms of Goffman's (1974) theory of frames. The social movements literature has relied heavily on the conception of "mobilization frames" (see Snow et al., 1986; Benford and Snow, 2000). With a strong focus on frame "resonance," collective action framing theory suggests that social movements' major problem in mobilization is the construction of frames that are likely to motivate people to join or to continue participation. That said, the mobilization framing literature tends to treat goals as constant, and to focus on *how* work should be done (e.g., through discussion of a "revolutionary frame"), rather than on the role that frames play as goals or targets.

The broader sociological literature provides us with several other tools to think about how clear definitions of genres might impact the success of organizations. Zuckerman (1999) uses firm performance data to argue that a lack of conformity to categories – measured by under-coverage by

stock analysts – can result in undervaluation of publicly traded firms. In the innovation literature, Hargadon and Douglas (2001) suggests that new innovations face barriers to diffusion when users do not understand how to use them. Hargadon and Douglas describe how Edison was able to sell electrification only after he re-framed it as a safer form of gas lighting through steps that included reducing the brightness of lights to more closely match what consumers expected from gas.

Work by scholars in information systems and organizations have explored another category-based effect in the adoption and effective use of new technological products. Building on Goffman (1974), Orlikowski (1992) and Orlikowski and Gash (1994) show that collaborative software might be adopted, but used with only a subset of its features or abilities, if users do not approach the new tool with a cognitive frame that can make sense of the new functionality. In order to take advantage of new modes of work within an organization, users must understand what that new type of work could entail. In this research, users' inability to frame the nature of the processes at play is associated with the failure of new technologies. Although they use a different success metric, Schweik and English (2012) find that the strongest predictor of early-stage success for FLOSS projects is the number of categories that projects are labeled with in their database. Schweik and English are surprised by this result and attribute it, tentatively, to strong leadership. An alternative explanation is the number of categories reflects a framing effect of the type I propose.

*Proposition 2: Low Barriers to Contribution*

Many theories of Wikipedia's success have emphasized the low barriers associated with contribution. Benkler's (2006) theory of peer production is based, fundamentally, around the idea that Internet-mediated communication has lowered transaction costs enormously. This model suggests that many people have a latent motivation to participate in the provision of public goods on the Internet that is higher than the direct costs of doing so, but that this motivation is frequently still too low to offset barriers in the form of transaction costs associated with contributing through contracting on a market or creating or joining a firm.

For example, if a hypothetical user finds a small error while reading an encyclopedia article – e.g., a missing comma – she may be motivated to spend a small amount of time and effort to fix the



error and share that modification. That said, in order to contribute, her motivation would have to be high enough to offset the costs of making that change. If the user is reading Encyclopedia Britannica without the Internet, making the change would involve finding the address of the publisher, finding a letter and envelope, writing a letter notifying the editors of the missing comma and noting the page and edition, paying for a stamp, and sending the letter. The entire experience might take half an hour and cost US\$1 in materials. Although many people might notice the error and be motivated to fix it, fewer users would be motivated enough to invest the time and money to follow through. If, on the other hand, there is an *Edit* button at the top of the page that the user is reading, the act of fixing the comma for others may have a cost much closer to zero, and the comma will be fixed more often.

Alluding to this dynamic, many OCEP initiators suggested that Wikipedia mobilized large numbers of volunteers because their OCEPs introduced what, with foresight, they consider relatively unimportant barriers to contribution. Several initiators suggested that Wikipedia succeeded because it made contributing easy – especially for lightly motivated potential contributors. Frequently cited examples of barriers to contribution included account creation, the requirement to log in to contribute, the need to learn markup or simple codes to format text and to create hyperlinks, and the requirement to preview changes before submitting them.

For example, an initiator of h2g2 explained that he felt that Wikipedia was easier to edit than his project in part because it did not require an account or any longer-term engagement with the project:

It's... I think definitely the immediacy of it. And certainly one of the aspects was the fact that you don't have to sign up to edit. That you can look at a page and see something wrong and immediately edit without having to do anything else. Y'know? You can come along and do a drive-by-edit and never be involved again, and make a contribution. You can't do a drive-by on almost any other project (Interview).

Many initiatives voiced a similar sentiment. An initiator of GNE suggested that Wikipedia had succeeded where GNE failed because, "Wikipedia was so much simpler" (Interview) and explained that

– both socially as well as technologically – Wikipedia did more to reduce barriers to contribution.

Several initiators referred to the fact that contributing to Wikipedia required relatively little or no association with the project and pointed to the absence of a steep “learning curve” which might hinder joining. Indeed, the idea that even relatively small barriers can prevent the adoption of a technology-based product is a fundamental assumption behind the academic and practitioner focused literatures on usability (e.g., Nielsen, 1993; Preece, 2000). Several project initiators suggested that contributions by previously uninvolved contributors to Wikipedia were easier relative to other OCEPs.

While explaining why his project failed to take off while Wikipedia did not, the initiator of TDEP suggested that costs associated with contributing to the project were the major reason that TDEP failed where Wikipedia succeeded only several years later. In the case of TDEP, barriers included learning technology like the HyperText Markup Language (HTML) in order to contribute. Although HTML is widely known, and although contributing to Wikipedia also requires learning its own idiosyncratic markup, Wikipedia’s markup (called “wiki text”) offers a lower barrier to contribution because it looks less like “code” and more like plain text. As a result, people that do not know wikitext can still make workable, if imperfect, contributions (Leuf and Cunningham, 2001).

On the other hand, many initiators expressed skepticism that barriers to contribution reflected a full explanation for Wikipedia’s success. For example, only one initiator (TDEP) cited barriers to contribution as the single most important reason for Wikipedia’s success relatively to their OCEP. Additionally, several initiators argued that their OCEP was effectively as easy to contribute to as Wikipedia. For example, Aaron Swartz pointed to several small barriers associated with contributing to TheInfo:

So, one problem was a mandatory preview step before you saved it. Which probably wasn’t enough to kill the site single handedly, but I probably would have changed it [if given the opportunity to the project again]. Wikipedia does fine without that (Interview).

Swartz, like several other initiators, suggested that Wikipedia's success without certain barriers might have altered their own designs. But many felt that Wikipedia's barriers were still significant and that learning to use a computer, being comfortable with wiki text, and navigating Wikipedia's complex social organization, were impediments on the level of barriers in their systems. To this day, the Wikimedia foundation cites usability issues as a major barrier to contribution in Wikipedia.

Although initiators were able to point to barriers to contribution in their systems that were higher than in Wikipedia, they also pointed to many technical similarities and to relatively lower barriers in their OCEPs. Wikipedia launched with off-the-shelf software that imposed many limitations relative to other OCEPs. Finally, and although the determination is inherently subjective, I felt that, in 2001, several OCEPs appeared and operated more like Wikipedia does today than Wikipedia did when it launched.

### *Proposition 3: Low Social Ownership*

In the literature on peer production, reputation is frequently cited as an incentive for contribution (Lerner and Tirole, 2002; Benkler, 2006, 2013). This theory suggests that individuals are paid in reputation or status in the absence of monetary payments. Toward this end, many creators of peer production and social computing systems have designed systems that provide contributors with reputation and status through identity systems that attribute contributions to their authors (Monroy-Hernández et al., 2011; Kraut and Resnick, 2012). Although initiators of OCEPs often took these ideas for granted in the design of their own systems, several pointed to the fact that Wikipedia's own design and technology provided very little indication of authorship. This, they suggested, led to low degrees of territoriality (Brown et al., 2005) and social ownership of content. This, initiators argued, helped Wikipedia mobilize contributors relative to their efforts.

Early work on incentives in peer production was influenced by economics and treated FLOSS's ability to mobilize large communities in the absence of monetary rewards to be confounding. Several early FLOSS studies looked to reputation as a form of alternative currency. For example, Lerner and Tirole (2002) suggested that people contribute to FLOSS because they can build their reputations by doing so and they will be able to translate this reputation into higher human capital and

higher paying jobs. Reputation can also be internal to communities where one's contributions to a public good are rewarded through the increased opinion, and subsequent actions, of other members of the community (Willer, 2009b,a). This approach to reputation systems has been cited more broadly as an important driver of contributions in online communities and social computing systems (Dellarocas, 2006, 2010). For examples, FLOSS projects hosted on SourceForge have administrators, developers, junior developers, and other categories of contributors.

A pre-requisite to reputation building, of course, is a way for users to attach their identity to their contributions. Many FLOSS systems reflect reputation through formal roles within projects. However, when one visits a Wikipedia article, it is unclear who has produced it (Viégas et al., 2007). Several OCEP initiators cited Wikipedias' absence of a system for easily determining who did what as a reason for Wikipedia's success. Counterintuitively, they argue that Wikipedia attracted contributors because low attribution facilitated less individual social ownership of work products and less socially risky collaboration. Most failed projects used stronger attribution which led to more "territoriality" (Brown et al., 2005). Many OCEPs allowed little or no direct collaboration on text (i.e., collaborators would need to copy and replace text to improve it). Most of the projects that did allow some direct collaboration still had explicit authorship norms (see Jaszi, 1991; Coombe, 1998) which identified an individual or group as the "author" of any particular piece of content.

The Interpedia was one project that did not allow for collaboration at the level of article text. Interpedia's design called for individuals to work together to build an encyclopedia, but to have each article produced individually. For example, one initiator of Interpedia suggested:

People would contribute articles and then they would be missing or they'd be somewhere down the list below the default article. But they'd be the work of one identifiable person. Unlike the Wikipedia thing where it'd be very difficult to track down who contributed some messages (Interview).

A different initiator of Interpedia made a similar point:

But the Wikipedia sort of conquered because anyone could just write anything on any

page without anyone's approval (Interview).

In this way, the design of several OCEPs called for writing articles that were the identifiable work of a single individual. Of course, all content in all OCEPs was licensed permissively so that anyone could modify and build on any piece of contributed content. But for projects like Interpedia and TDEP, contributors were expected to ask the original author of each article to update the text, or to write their own version of the article by incorporating their work into a significantly different article to which they could claim authorship.

Another project that did not allow direct collaboration on text was E2. In E2, each topic or "node" could have a number of different "write-ups" or articles associated with it. Each of these write-ups was listed as being produced by the author who started the write-up. Although it was possible for some users to edit others' articles, the fact that each write-up was labeled as having been written by an identifiable author led to a social norm that looked down upon this type of intra-textual collaboration. If an E2 user saw something wrong with an article, they were encouraged to provide an improved write-up of their own rather than trying to edit somebody else's. E2 initiator Nate Oostendorp explained that the clear presence of authorship led to failures to mobilize by comparing their design to Wikipedia's:

In Wikipedia, when you submit content, you don't really get an authorship credit directly. Y'know, you appear in the history but these aren't necessarily your words, they are just sort of your contribution to Wikipedia. But with Everything ... their writings were still theirs. They had control of them on the site. And they received direct attribution. I think there was some weakness there in that when people wrote something and if it was factual content and if they had information that was incorrect, there was no real... I mean, occasionally an editor would go in and change the content but otherwise it was sort of up to them to receive communication and re-add to it (Interview).

Indeed, authorship and credit were cited repeatedly by initiators as more important than technical workflow limitations. Another early contributor and moderator of E2, Cliff Lampe, expounded on

this observation:

So I think having one article as opposed to several write-ups on a node took advantage of marginal contributions in a way that E2 is not set up to. That really helped make it a much more strongly “many hands make lighter work” type of exercise (Interview).

In the projects in my sample, ownership over text was often purely social. Several OCEPs whose initiators suggested that textual ownership hindered contribution allowed, from a technical perspective, for at least some direct collaboration on the text. But because they also listed authorship explicitly, a sense of social ownership over text meant this technical possibility was rarely taken advantage of.

Of course, Wikipedia is not immune to social ownership of text. For example, Thom-Santelli et al. (2009) have described territoriality in Wikipedia. And although Thom-Santelli et al. suggest that territoriality has benefits, they point to drawbacks within Wikipedia as well. By placing a name on the top of the page, many OCEPs made it socially risky to engage in direct collaboration on text. Although counterintuitive, this result finds some support in recent research in social computing (Bernstein et al., 2011); in some cases, users can engage and collaborate *because* there are not clearly identifiable “toes” to be stepped on. When these toes are visible, many potential collaborators chose not to build on each others’ work. Ironically, the fact that Wikipedia made authorship less scrutable opened the door to deeper and more widespread collaboration.

### *Synthesis*

Results of data analysis suggest three proposition to explain Wikipedia’s relative success at mobilizing volunteers: (P1) a reliance on a product or goal familiar to potential contributors; (P2) low barriers to contribution; and (P3) low social ownership of content. These results are summarized for each project in dichotomous codes shown in in Table 2. Several patterns can be seen in these results. First, Wikipedia is the only project coded as having satisfied each proposition. Second, P2 and P3 are largely coincident in the dataset – i.e., only two projects (E2 and h2g2) have different codes for P2 and P3. And although they are conservatively coded “No” for P3, both h2g2 and Ev-

Project	P1: Familiar Goal	P2: Low Barriers	P3: Low Social Ownership
Interpedia	No	No	No
TDEP	Yes	No	No
GNE	Yes	No	No
Everything2	No	Yes	No
h2g2	No	Yes	No
Nupedia	Yes	No	No
TheInfo	No	Yes	Yes
Wikipedia	Yes	Yes	Yes

Table 2: Dichotomous codes for each encyclopedia project for each of the three propositions described in the result. A code of “Yes” suggests that there was strong support for that theme in the data associated with the project while “No” suggests there was not strong support.

		Innovativeness of Goal/Product	
		Familiar (P1=Y)	Novel (P1=N)
Innovativeness of Process/Tools Novel (P2,3=Y) Familiar (P2,3=N)	Familiar (P2,3=N)	TDEP GNE Nupedia	Interpedia  Everything2
	Novel (P2,3=Y)	Wikipedia	h2g2  The Info Network

Figure 2: Two-by-two table adapted from Figure 1. Propositions are added to the labels on the axes and the names of OCEPs are placed onto the grid based on their coding as described in Table 2.

everything *did* allow some editing of each others content. It is also the case that no project other than Wikipedia is coded “Yes” for P1 is also coded “Yes” for *either* P2 or P3.

This pattern of results can be understood more clearly if the propositions are compared with traditional encyclopedias. Toward that end, we can return to the diagram in Figure 1. Figure 2 reproduces that diagram but omits descriptions of each cell and, instead, uses the codes in Table 2 to place each OCEP onto the grid. In the figure, P1 is mapped onto the x-axis and reflects the novelty of the product or goal. Projects in the left column are the OCEPs that hewed closely to existing def-

initions of encyclopedias in terms of their goal and scope. In the right column are projects that tried to expand upon this concept. In the context of traditional encyclopedias, P2 and P3 each reflects innovation in the process and tools used to organize production. This reflects, for example, what Benkler (2006) suggests is Wikipedia's, "radically new form of encyclopedia writing." Traditional encyclopedia writing was marked by intentional barriers to participation by non-experts and by rigid systems of workflow, roles, and processes designed to ensure quality. Because P2 and P3 were largely coincident in the dataset, I have mapped them, in combination, to the y-axis in the Figure 2. Projects in the top row aimed to organize production, authority, and social ownership in ways that tracked traditional methods. Projects in the bottom row used innovative processes and sought to remove barriers to contribution and social ownership.

TDEP, GNE, and Nupedia are coded as residing in the top-left quadrant. These three projects sought to build traditional encyclopedias using traditional processes and methods. They differed from more traditional encyclopedias like Britannica primarily in that they released their content freely on the Internet and did not pay authors. Each of these projects struggled to mobilize contributors and to compete with existing encyclopedias. In the bottom-right quadrant is TheInfo which reflects the only project coded as attempting to expand upon traditional encyclopedic frames while discarding traditional methods of production. Despite press and attention, TheInfo struggled to attract a community. Projects along this diagonal (top-left to bottom-right) were the least successful efforts in the sample in terms of their ability to mobilize contributors. In general, efforts that used a mix of innovation and familiarity worked better than either very innovative or very familiar projects.

Projects on the other diagonal (top-right to bottom-left) are coded as mixing innovation with tradition. Would be contributors to projects on this diagonal were likely to be comfortable and familiar with either the goal or the process at the core of the project, but not with both. Three of the four projects along this diagonal are the projects that built sustainable communities and persist today. Interpedia is coded as falling into the top left quadrant. E2 and h2g2 are coded as having low barriers to entry but high level of social ownership over work products. Because of their mixed coding on P2 and P3, they are placed on the border of the two cells in the right column. These three



projects sought to build innovative genres of reference works, but did so using methods that were at least moderately similar to traditional models. In the bottom-left shaded corner is Wikipedia, the only projects that sought to build a traditional encyclopedia using novel methods.

Taken together, these propositions suggest that Wikipedia succeeded in building a large community of volunteer contributors because it sought to build a product that potential contributors were already familiar with, but aimed to do so in a way that used novel peer productive forms of organizing volunteer labor that took advantage of advances in communication technology. Although Wikipedia's initiators encouraged editors to cast aside preconceptions of *how* an encyclopedia should be written, they appealed explicitly to their contributors' existing conceptions of what an encyclopedia should be. My results suggest that Wikipedia succeeded because it sought to do the old thing in a new way.

Consideration of other peer production projects can find anecdotal support for this dynamic. Many of the largest and most successful peer production communities would, if placed on the grid, fall unquestionably into the shaded quadrant. GNU/Linux, frequently referred to simply as the Linux operating system, began as a project to reproduce UNIX. Even the name "GNU" stands for "GNU's Not Unix" – a joke because GNU's not UNIX in name only. The popular OpenOffice and LibreOffice office suites are free software re-implementations of Microsoft Office. The GCC compiler and the Apache webserver each implement published standards where the project's stated goal was literally written down in a standards document before any contributor arrived. OpenStreetMap seeks to create a free replacement to Google Maps.

## ALTERNATIVE EXPLANATIONS

Although the inductive approach used in this paper can produce new theory, it cannot test these theories and it cannot reject the possibility that there are other valid explanations that did not emerge from my data through my methods. That said, there are two commonly cited explanations for Wikipedia's success that found disconfirming evidence in my data and analysis. First, there is little support in my dataset for the argument that (AE1) Wikipedia succeeded because it was technologically superior to alternatives. On the contrary, several project initiator cited the relative techno-

logical unsophistication of Wikipedia as an indirect reason for its success. Second, there was little support for the argument that (AE2) Wikipedia succeeded where other projects fell short because Wikipedia's timing was correct and that other projects were simply ahead of their time.

*Alternative Explanation 1: Technological Superiority*

If not a technological advance itself, Wikipedia is frequently described as the product of a technological inevitability. Evidence in my data set provides little support for the argument that Wikipedia succeeded because of its technological superiority. By definition, every OCEP used the Internet. Indeed, most OCEPs in my sample used technology at least as sophisticated as Wikipedia's and most used technology that was more sophisticated. For example, an initiator of Interpedia – a project that predated Wikipedia by nearly a decade – explained:

A lot of the stuff in the Wikipedia is extremely obvious and not very sophisticated. I mean, the Wikipedia is not high tech. I always imagined something high tech. That's my nature. I envision things that are of a higher technical level. We envisioned for the Interpedia as something that would be high tech. And we could see the Interpedia inspiring the Wikipedia, but not the other way around (Interview).

On the other hand, many other projects had a strong focus on technological development. For example, TheInfo's initiator Aaron Swartz explained he saw his own role as a technological facilitator saying:

I had this notion that my job was to provide the platform ... And so, I kept trying to refine the user interface and things like that to make it more inviting so more people would write stuff (Interview).

Similarly, contributors to GNE worked on code over many months and produced several competing technological implementations of software to run the system that each went through multiple releases.

On the other hand, Wikipedia's initiators were, by far, the least technically sophisticated founding team in my sample. Every other initiator, with the exception of Nupedia which shared a set of founders with Wikipedia, was a technologist. In each case, initiators thought of their projects in primarily technological terms. Despite the fact that none of the questions in my interview protocol explicitly mentioned technology or software, a large portion of many interviews focused on issues of the design of the technological systems that facilitated interaction.

Wikipedia is the one OCEP in my sample that did not begin by writing its own software. Instead, it used an off-the-shelf, freely available, piece of wiki software called *UseModWiki*.<sup>7</sup> As non-technologists initially more focused on Nupedia, Wikipedia's founders invested very few resources into technology for the project. Two OCEP initiators suggested that Wikipedia may have succeeded due to this difference in focus and their own myopic focus on technology. While a strong technological orientation was rarely cited as hurting OCEPs, it may have distracted from the key social and organizational problems of building a community of contributors.

#### *Alternative Explanation 2: Timing*

Another commonly cited alternative explanation for the success of Wikipedia relative to other OCEPs is that Wikipedia simply had better timing. Each of the seven projects in my sample was founded before Wikipedia. As a result, one explanation for Wikipedia's success was that building an OCEP was an idea whose time came in early 2001 and that earlier projects were simply ahead of the curve. Had they been started when Wikipedia was, this alternative explanation suggests, other OCEPs might have succeeded instead. Support for this perspective come from the fact that Internet use was expanding rapidly around 2001 with millions of new Internet users being added each month.<sup>8</sup> Additionally, Internet technology was being developed rapidly around the turn of the millennium. OCEPs founded before 2001 may have simply been at a disadvantage relatively

---

<sup>7</sup>A good resource with details on UseModWiki is the Wikipedia article on the subject: <http://en.wikipedia.org/wiki/UseModWiki> (Accessed June 29, 2013)

<sup>8</sup>The International Telecommunications Union publishes data on Internet penetration rates globally. Details and downloadable data are available online at <http://www.itu.int/en/ITU-D/Statistics/Pages/stat/default.aspx> (Accessed June 29, 2013). However, there is very little good data available on Internet penetration in the mid 1990s. Data from the Internet Systems Consortium suggests that connected hosts increased from several hundred thousand hosts to several tens of millions – an increase of over 100 times over the 1990s: <https://oldmirage.isc.org/solutions/survey/history> (Accessed June 29, 2013).

to Wikipedia in that they were mobilizing from a smaller pool of potential contributors and using less sophisticated tools. Although I found some support for this explanation in one case, my data suggests that timing was not a critical factor in explaining the relative failure of most OCEPs in my sample.

Interpedia was launched in 1993 when the web was still in its infancy. One of the initiators of Interpedia reflected on Interpedia's failure saying:

It could have been timing. That was part of it. We didn't have the mechanism of the web which is clearly a lot better way. Not just hypertext, but the web, as it's become so large. And that's another thing. We didn't have the installed user-base – for want of a better term – of people that have time. And so, [for Wikipedia], the timing was right (Interview).

Citing both the state of technological development and the state of potential contributor-bases, both Interpedia initiators that I interviewed alluded to timing as a possible contributor to their own projects struggles to mobilize relative to Wikipedia. Indeed, Interpedia was founded nearly one decade before Wikipedia at a point when the Internet was much smaller than it was in 2001.

But Interpedia was the only project that mentioned timing as a reason for their effort's failure to mobilize volunteers. All of the other OCEPs in my sample launched within 4 years of Wikipedia's founding and were active less than two years before Wikipedia launched. This includes TDEP which was the only other project that ended before Wikipedia launched. With the exception of Interpedia and TDEP, every project was launched within three years of Wikipedia; three were launched less than one year before, and one project launched less than a week before Wikipedia in January 2001. A plot of the periods during which each project was active is shown in Figure 3. In five of the seven early OCEPs, efforts continued in the period before, during, and after Wikipedia's launch. Moreover, Wikipedia grew slowly over its first year. Several of the OCEPs in my sample – most notably E2 and h2g2 – were large and more established communities into 2002 or 2003.

It was also not the case that Wikipedia made use of technology unavailable to other projects. As detailed in Section AE1 above, Wikipedia launched using off-the-shelf software, available under a

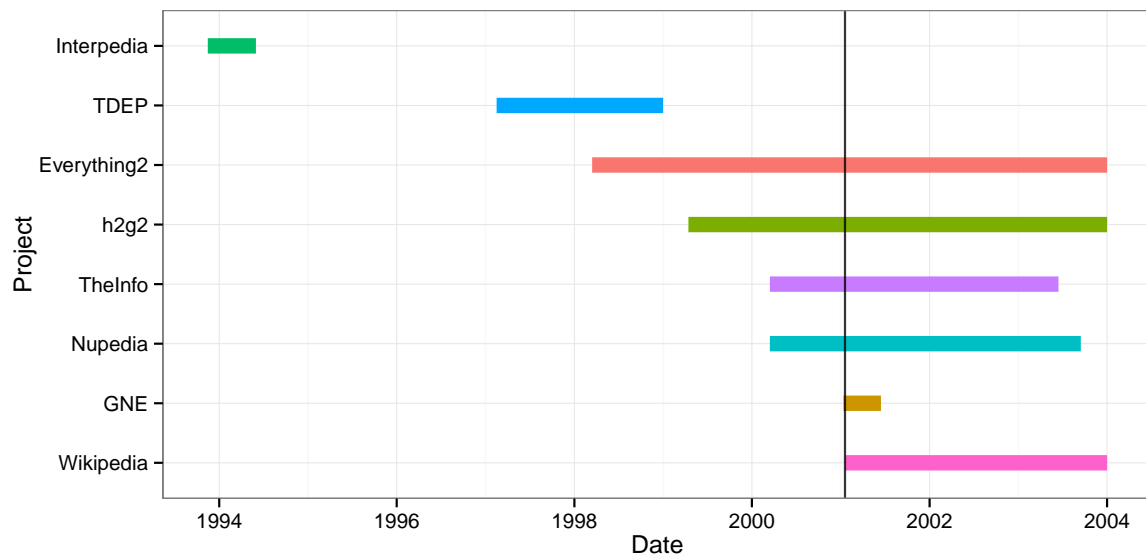


Figure 3: Gantt chart showing when OCEPs were active. Everything2, h2g2, and Wikipedia remain active today. The vertical line marks January 18, 2001 – the date that Wikipedia was publicly launched.

free software license. The software Wikipedia used was released in 1999 and was written in Perl – the same programming language used in several of the other OCEPs. As a result, Wikipedia’s exact technology would have been easy to use, customize, and build upon for other OCEP initiators. Indeed, archival data from TheInfo shows that a contributor suggested that the site administrators consider using wiki software saying, “there’s a thing called WikiWiki that you should go check out” (Archival Data). Initiator Aaron Swartz rejected the suggestion, preferring the technological affordances of his own system, explaining later that, “I was frustrated with [wiki software]. And you can see that I fixed a lot of things that ended up being in [later version of] Wikipedia” (Interview). Although timing may have played a role in the earliest OCEPs in my sample, it was rejected as an explanation by most projects initiators and found little support in my analysis.

## CONCLUSION

Previous scholarship has attributed Wikipedia’s success in mobilizing to an accident or simply to luck (Reagle, 2009). But what is luck? Luck, according to one perspective, is simply a term we invoke when we do not understand and cannot control the processes that determine some fortunate

outcome. Through this work, I have attempted to deconstruct and explain Wikipedia's "luck" to offer a theory of volunteer mobilization in peer production. I have suggested three propositions for why Wikipedia succeeded in mobilizing a large community of volunteer contributors where very similar efforts failed. I have suggested that Wikipedia succeeded because it sought to build a product that was familiar to potential contributors while adopting innovative processes and methods of organizing labor. This pattern, I suggest, is common to many of the most successful examples of peer production.

There are limitations to the methodology employed in this essay and my results are only propositions. In future work, I hope to operationalize these propositions and test them as hypotheses in quantitative analyses. There are several large databases of FLOSS projects like SourceForge, large datasets of populations of wikis (e.g., Kittur and Kraut, 2010; Reich et al., 2012), and remixing communities (e.g., Hill and Monroy-Hernández, 2013), which, if properly coded in terms of these propositions, may support empirical tests of the relationships suggested.

But if luck refers to a lack of control, we must also acknowledge that even successful tests of these propositions may not mean that we can affect mobilization. For example, there remain important open questions about the generalizability of these findings. The results in this paper do not tell us, *prima facie*, if the particular innovative methods of organizing labor that Wikipedia used are likely to work for other types of creative goods or in projects with very different goals. Additionally, these results suggest limits in terms of what nascent projects set out to do. Many efforts that might seek to learn from these results to increase the likelihood of their own success will begin with goals in mind. Advice to "do something else," or to "try to build a different type of product," may be neither actionable nor useful.

These results may be both valid and limited in predictive power. In experiments around music popularity, Salganik et al. (2006) have shown that it is effectively impossible to predict the success of songs by looking only at qualities of the songs themselves because popularity is largely driven by social interactions that are not caused by qualities of the music. Predicting the success of firms or products seems similar, at least superficially, to the goals of this study and it is notoriously difficult to build strong predictions (e.g., Watts, 2011). Quantitative tests of these hypotheses might fully

confirm that projects with a familiar product and innovative process are more likely to mobilize volunteers than products that are innovative on both dimensions – but the effect size, in absolute terms, might be small. Although the proposed relationships might hold, on average, projects that “do everything right” may still be overwhelmingly likely to fail for reasons outside of initiators’ power.

With these limitations foregrounded, this paper offers an important set of contributions to the literature on peer production. From a methodological perspective, this paper offers an answer to the critique that the peer production literature has spent too much time studying successful projects. If peer production scholars care about mobilization, this paper reflects a step forward in that it stops selecting on the dependent variable and, through that process, brings relative failures into the analysis. More importantly, it offers a theory about why some attempts to build peer production projects lead to large mobilizations, like Wikipedia and Linux, while the vast majority struggle to attract even a second contributor. It offers a step, supported in empirical evidence, toward a broader causal theory of mobilization.

## REFERENCES

- Benford, R. D. and Snow, D. A. (2000). Framing processes and social movements: An overview and assessment. *Annual Review of Sociology*, 26(1):611–639.
- Benkler, Y. (2002). Coase’s penguin, or, linux and the nature of the firm. *Yale Law Journal*, 112(3):369.
- Benkler, Y. (2006). *The Wealth of Networks: How Social Production Transforms Markets and Freedom*. Yale University Press.
- Benkler, Y. (2013). Peer production and cooperation. In Bauer, J. M. and Latzer, M., editors, *Handbook on the Economics of the Internet*. Edward Elgar.
- Bernstein, M., Monroy-Hernández, A., Harry, D., André, P., Panovich, K., and Vargas, G. (2011). 4chan and /b/: An analysis of anonymity and ephemerality in a large online community. In *Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media*, pages 50–57, Barcelona, Spain. AAAI Press.
- Berzroukov, N. (1998). A second look at the cathedral and the bazaar. *First Monday*, 4(12).
- Brown, G., Lawrence, T. B., and Robinson, S. L. (2005). Territoriality in organizations. *The Academy of Management Review*, 30(3):577–594.

- Bush, V. (1945). As we may think. *Atlantic Monthly*, 176(1):101–108.
- Charmaz, K. (2006). *Constructing Grounded Theory: A Practical Guide through Qualitative Analysis*. Sage Publications Ltd, 1 edition.
- Coase, R. H. (1937). The nature of the firm. *Economica*, 4(16):386–405.
- Coombe, R. J. (1998). *The Cultural Life of Intellectual Properties: Authorship, Appropriation, and the Law*. Duke University Press Books.
- Corbin, J. M. and Strauss, A. C. (2007). *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory*. Sage Publications, Inc, 3rd edition.
- Crowston, K., Annabi, H., and Howison, J. (2003). Defining open source software project success. In *Proceedings of International Conference on Information Systems*.
- Crowston, K., Wei, K., Howison, J., and Wiggins, A. (2010). Free/libre open source software: What we know and what we do not know. *ACM Computing Surveys*, 44(2):2012. Working Paper.
- Dellarocas, C. (2006). Reputation mechanisms. In *Handbook on Economics and Information Systems*, page 2006. Elsevier Publishing.
- Dellarocas, C. (2010). Designing reputation systems for the social web. *SSRN eLibrary*.
- Eisenhardt, K. M. (1989). Building theories from case study research. *The Academy of Management Review*, 14(4):532–550.
- English, R. and Schweik, C. M. (2007). Identifying success and tragedy of FLOSS commons: A preliminary classification of sourceforge.net projects. In *Proceedings of the First International Workshop on Emerging Trends in FLOSS Research and Development*, page 11. IEEE Computer Society.
- Goffman, E. (1974). *Frame Analysis: An Essay on the Organization of Experience*. Harvard University Press, first edition edition.
- Hargadon, A. B. and Douglas, Y. (2001). When innovations meet institutions: Edison and the design of the electric light. *Administrative Science Quarterly*, 46(3):476–501.
- Healy, K. and Schussman, A. (2003). The ecology of open-source software development.
- Hill, B. M. (2005). Reflections on free software past and future. *First Monday*, 10(10). Special issue.
- Hill, B. M. and Monroy-Hernández, A. (2013). The remixing dilemma the trade-off between generativity and originality. *American Behavioral Scientist*, 57(5):643–663.
- Howe, J. (2008). *Crowdsourcing: Why the power of the crowd Is driving the future of business*. Crown Business, 1 edition.
- Huang, R. (2012). RQDA: r-based qualitative data analysis. R package version 0.2-3.



- Jaszi, P. (1991). Toward a theory of copyright: The metamorphoses of "Authorship". *Duke Law Journal*, 1991(2):455.
- Kittur, A. and Kraut, R. E. (2010). Beyond wikipedia: coordination and conflict in online production groups. In *Proceedings of the 2010 ACM conference on Computer supported cooperative work*, pages 215–224, Savannah, Georgia, USA. ACM.
- Kraut, R. E. and Resnick, P. (2012). *Building Successful Online Communities: Evidence-Based Social Design*. The MIT Press.
- Krishnamurthy, S. (2002). Cave or community? an emperical examination of 100 mature open source projects. *First Monday*, 7(10).
- Krishnamurthy, S. (2005). The elephant and the blind men - deciphering the Free/Libre/Open source puzzle. *First Monday*, 10(10). Special issue.
- Lakhani, K. R. and von Hippel, E. (2003). How open source software works: "Free" user-to-user assistance. *Research Policy*, 32(6):923–943.
- Lerner, J. and Tirole, J. (2002). Some simple economics of open source. *Journal of Industrial Economics*, 50(2):197–234.
- Leuf, B. and Cunningham, W. (2001). *The Wiki way: quick collaboration on the Web*. Addison-Wesley Longman Publishing Co., Inc. Boston, MA, USA.
- Lih, A. (2009). *Wikipedia Revolution, The: How a Bunch of Nobodies Created the World's Greatest Encyclopedia*. Hyperion, first edition. first printing. edition.
- Lin, Y. (2005). The future of sociology of FLOSS. *First Monday*, 10(10). Special issue.
- Loveland, J. and Reagle, J. (2013). Wikipedia and encyclopedic production. *New Media & Society*.
- Luther, K., Caine, K., Zeigler, K., and Bruckman, A. (2010). Why it works (when it works): Success factors in online creative collaboration. In *Proceedings of the ACM Conference on Supporting Group Work*, Sanibel Island, Florida, USA. ACM.
- MacCormack, A., Rusnak, J., and Baldwin, C. Y. (2006). Exploring the structure of complex software designs: An empirical study of open source and proprietary code. *Management Science*, 52(7):1015–1030.
- MacCormack, A. D., Rusnak, J., and Baldwin, C. Y. (2012). Exploring the duality between product and organizational architectures: A test of the "mirroring" hypothesis. *Research Policy*, 41(8):1309–1324.
- McCarthy, J. D. and Zald, M. N. (1977). Resource mobilization and social movements: A partial theory. *The American Journal of Sociology*, 82(6):1212–1241.
- Minkoff, D. C. (1999). Bending with the wind: Strategic change and adaptation by women's and racial minority organizations. *The American Journal of Sociology*, 104(6):1666–1703.

- Mockus, A., Fielding, R. T., and Herbsleb, J. D. (2002). Two case studies of open source software development: Apache and mozilla. *ACM Transactions on Software Engineering Methodology*, 11(3):309–346.
- Monroy-Hernández, A., Hill, B. M., Gonzalez-Rivero, J., and boyd, d. (2011). Computers can't give credit: how automatic attribution falls short in an online remixing community. In *Proceedings of the 2011 annual conference on Human factors in computing systems*, pages 3421–3430, Vancouver, BC, Canada. ACM.
- Murray, F. and O'Mahony, S. (2007). Exploring the foundations of cumulative innovation: Implications for organization science. *Organization Science*, 18(6):1006–1021.
- Nelson, T. H. (1981). *Literary Machines*. Self Published, Swarthmore, PA, 3rd edition.
- Nielsen, J. (1993). *Usability Engineering*. Academic Press Inc.
- Oostendorp, N. (1999). Everything.
- Orlikowski, W. J. (1992). Learning from notes: organizational issues in groupware implementation. In *Proceedings of the 1992 ACM conference on Computer-supported cooperative work*, pages 362–369, Toronto, Ontario, Canada. ACM.
- Orlikowski, W. J. and Gash, D. C. (1994). Technological frames: making sense of information technology in organizations. *ACM Transactions of Information Systems*, 12(2):174–207.
- Ortega, F. (2009). *Wikipedia: A Quantitative Analysis*. PhD dissertation, Universidad Rey Juan Carlos.
- Postma, R. (2000). Everything FAQ.
- Preece, J. (2000). *Online Communities: Designing Usability and Supporting Sociability*. Wiley, 1 edition.
- Raymond, E. S. (1999). *The cathedral and the bazaar: Musings on Linux and open source by an accidental revolutionary*. O'Reilly and Associates, Sebastopol, CA.
- Reagle, J. (2009). Wikipedia: the happy accident. *interactions*, 16(3):42–45.
- Reagle, J. M. (2008). *In Good Faith: Wikipedia Collaboration and the Pursuit of the Universal Encyclopedia*. PhD, New York University.
- Reich, J., Murnane, R., and Willett, J. (2012). The state of wiki usage in U.S. k–12 schools leveraging web 2.0 data warehouses to assess quality and equity in online learning environments. *Educational Researcher*, 41(1):7–15.
- Salganik, M. J., Dodds, P. S., and Watts, D. J. (2006). Experimental study of inequality and unpredictability in an artificial cultural market. *Science*, 311(5762):854–856.
- Schweik, C. and English, R. (2012). *Internet Success: A Study of Open Source Software Commons*. MIT Press, Cambridge, Massachusetts.

- Shirky, C. (2010). *Cognitive Surplus: Creativity and Generosity in a Connected Age*. Penguin Press HC, The.
- Snow, D. A., Rochford, E. B., Worden, S. K., and Benford, R. D. (1986). Frame alignment processes, micromobilization, and movement participation. *American Sociological Review*, 51(4):464–481.
- Soule, S. A. and King, B. G. (2008). Competition and resource partitioning in three social movement industries. *The American Journal of Sociology*, 113(6):1568–1610.
- Stallman, R. M. (1999). Free universal encyclopedia and learning resource.
- Stallman, R. M. (2002). The GNU manifesto. In *Free Software, Free Society: Selected Essays of Richard M. Stallman*, page 224. Free Software Foundation.
- Tapscott, D. and Williams, A. D. (2008). *Wikinomics: How Mass Collaboration Changes Everything*. Portfolio Hardcover, expanded edition.
- Thom-Santelli, J., Cosley, D. R., and Gay, G. (2009). What’s mine is mine: Territoriality in collaborative authoring. In *Proceedings of the 27th international conference on Human factors in computing systems*, pages 1481–1484, Boston, MA, USA. ACM.
- Viégas, F. B., Wattenberg, M., McKeon, M. M., and Schuler, D. (2007). The hidden order of wikipedia. In *Online Communities and Social Computing*, HCII, page 445–454. Springer-Verlag, Berlin, Heidelberg.
- von Hippel, E. and von Krogh, G. (2003). Open source software and the ‘private-collective’ innovation model: Issues for organization science. *Organization Science*, 14(2):209–223.
- Watts, D. J. (2011). *Everything Is Obvious: How Common Sense Fails Us*. Crown Business.
- Wesler, H. T., Smith, M., Fisher, D., and Gleave, E. (2008). Distilling digital traces: Computational social science approaches to studying the internet. In *The SAGE Handbook of Online Research Methods*, page 116–140.
- West, J. and O’Mahony, S. (2005). Contrasting community building in sponsored and community founded open source projects. In *Proceedings of the 38th Annual Hawaii International Conference on System Sciences (HICSS’05) - Track 7 - Volume 07*, page 196.3. IEEE Computer Society.
- Willer, R. (2009a). Groups reward individual sacrifice: The status solution to the collective action problem. *American Sociological Review*, 74(1):23–43.
- Willer, R. (2009b). A status theory of collective action. In *Advances in Group Processes*, volume 26, pages 133–163. Emerald Group Publishing, Bingley.
- Williamson, O. E. (1981). The economics of organization: The transaction cost approach. *The American Journal of Sociology*, 87(3):548–577.
- Zuckerman, E. W. (1999). The categorical imperative: Securities analysts and the illegitimacy discount. *The American Journal of Sociology*, 104(5):1398–1438.

## APPENDIX A: INTERVIEW PROTOCOL

Thank you so much for taking the time to talk to me!

I'm a PhD student at MIT doing research on online encyclopedia projects and I am interested in [PROJECT NAME] which I know you were involved in starting. I've gone through some materials online that I dug up about [PROJECT NAME] and I'm interested in learning more and hearing the whole story from your perspective.

I've got a series of open-ended questions written down but you should feel welcome to take the conversation in directions you think are interesting. We can keep this pretty unstructured. Please let me know, at any point, if you have any questions for me and don't feel like you have to answer any questions you don't want to.

My plan is to write this up into a academic paper. If you'd prefer, I'll be happy to disguise your name and project in the published version and to keep my notes and recordings confidential. You can let me know later as well.

### *Background*

First, I'd love for you to tell me a little about yourself today – where you are and what you're up to?

Where do you live? (*Most of these interviews will be phone based.*)

If you don't mind me asking, how old are you now?

Can you tell me a little bit about your passions or hobbies?

Thinking back in time to when you created [PROJECT NAME], can you tell me about what you were up to then?

Where were you living?

How were you supporting yourself? Were you working?

If you don't mind my asking, how do you support yourself now?

How about hobbies or passions at the time?

### *Project Creation*

When did you first decide to create an online encyclopedia? What were the there projects or experiences that led up your decision to start [PROJECT NAME]?

Did you work with others to create [PROJECT NAME]? Who? What was the collaboration like?

Were you influenced by other projects, thinkers, or writers? Who/which ones? How?

Why did you start the project? What were the ideas or goals behind [PROJECT NAME]?

Can you describe the basic “design” of [PROJECT NAME]? How did it work? What did it look like?

Who did you imagine as the audience for the [PROJECT NAME]? What type of people were you trying to serve?

How did you try to market your project?

Do you think you were successful? Why? Or, why not?

Who did you imagine as the contributors to [PROJECT NAME]? What type of people did you imagine contributing?

How did you try to recruit contributors?

Do you think you were successful? Why? Or, why not?

What sort of resources were necessary (e.g., money, technical skills, labor, etc.)? What did you do to get these? Did you have trouble getting necessary resources?

At the time of your projects, there were other online collaborative encyclopedia projects. Did you know about these? What did/do you think of them?

[*Going through my list.*] Did you know about [OTHER PROJECT NAME]? What did you think of it? How was your project different? What did you think of them?

Are there other similar projects you think I've missed?

### *About Your Project*

Can you walk me through the history of the [PROJECT NAME]?

What were some of the greatest challenges you faced with [PROJECT NAME]?

What are some of your proudest achievements with [PROJECT NAME]? What did you think you did right, and why?

What are some of your greatest regrets? What do you think you did wrong, and why?

### *Your Project Today*

Where, if anywhere, is [PROJECT NAME] today?

If you could do your project over again, what would you do differently?

More generally, what has your experience with [PROJECT NAME] taught you? What have others learned?

Are you still in touch with any collaborators? What are they up to?

### *Wikipedia*

Superficially at least, Wikipedia seems to share a lot in common with [PROJECT NAME]. What do you think about Wikipedia?

How is Wikipedia similar, and different, to [PROJECT NAME]?

What has it been like watching Wikipedia become so enormously successful?

Have you contributed to Wikipedia? How do think the experience compared to work on [PROJECT NAME]? What was different?

Why do you think Wikipedia has worked out so much better than [PROJECT NAME]?

What do you think Wikipedia could learn from [PROJECT NAME]?

Do you think you've learned anything about [PROJECT NAME] from Wikipedia?

*Concluding*

Thank you *so much* for taking the time to talk to me! I appreciate all the information you've been able to give me.

If you have any questions for me, I'm happy to answer them. If you're curious about my research project or my hypotheses, I'm happy to give you details

If you know of other people I might want to talk to or other ideas of projects I might have missed, I'd love to hear that.