

UNRULY HORDES OR ALTRUISTIC COMMUNITIES?:
CROWDSOURCING IN ACADEMIC INSTITUTIONS

By

Kayla Utendorf

A Thesis Submitted in Partial Fulfilment
of the Requirements for the Degree of
Master of Arts in History

Middle Tennessee State University
May 2016

Thesis Committee:

Dr. Kelly Kolar, Chair

Dr. Susan Myers-Shirk

ACKNOWLEDGEMENTS

I would like to thank the members of my thesis committee, Dr. Kelly Kolar and Dr. Susan Myers-Shirk for their extensive help on this project. Dr. Myers-Shirk provided great advice on how begin research on a historical trend that is currently taking place. Dr. Kolar read more drafts than I care to count and pushed me not only to finish this project, but to finish with the strongest arguments and writing possible. I would also like to thank staff and volunteers at DIY History and Project Gutenberg Distributed Proofreaders for kindly filling out surveys, as well as staff at anonymous archives for agreeing to allow me to interview them on their crowdsourcing projects. They voiced the benefits of crowdsourcing better than I could ever have hoped to.

I would also like to thank my family and friends for putting up with all the canceled plans and excuses of “I have homework ... ” Lastly, I need to express the deepest gratitude to my fiancé, Steve, who has gladly dealt with all this and more as I spent every evening and weekend hunched over a computer. From turning down walks in the park to movie nights, he has stayed by my side, letting me focus when I needed to work and dragging me outdoors when I truly needed a break.

ABSTRACT

This thesis uses surveys, interviews, and blogs to examine crowdsourcing in archives, libraries, and other academic institutions, with special focus on the connection between crowdsourcing and traditional volunteering and on the benefits of crowdsourcing for archives. Traditional volunteers in libraries and archives and crowdsourcing volunteers are both motivated primarily by enjoyment in the task at hand and by a strong sense of community and friendship among their fellow volunteers. Crowdsourcing provides archives with an opportunity to achieve work that they would not have the resources to achieve otherwise and provides increased outreach opportunities by allowing volunteers to engage with archival records. This thesis also traces a history of crowdsourcing projects, in archives and otherwise, and provides a series of recommendations for those considering starting a crowdsourcing project.

TABLE OF CONTENTS

	Page
LIST OF FIGURES.....	vi
CHAPTER I: INTRODUCTION.....	1
LITERATURE REVIEW.....	2
METHODOLOGY.....	11
CHAPTER OUTLINE.....	12
CHAPTER II: MOTIVATIONS OF VOLUNTEERS.....	14
DEFINING “VOLUNTEERING”.....	14
MOTIVATION BEHIND FEDERAL VOLUNTEER PROGRAMS.....	18
MOTIVATIONS OF VOLUNTEERS.....	21
VOLUNTEERS IN LIBRARIES AND ARCHIVES.....	26
CONCLUSION.....	32
CHAPTER III: HISTORY OF CROWDSOURCING IN ACADEMIC INSTITUTIONS.....	34
DEFINING CROWDSOURCING.....	35
THE HISTORY OF INTERNET CROWDSOURCING.....	39
CROWDSOURCED PROJECTS.....	49
BUT WHAT HAPPENED TO THE GATEKEEPER?.....	58
CONCLUSION.....	68

CHAPTER IV: BENEFITS OF CROWDSOURCING IN ACADEMIC

INSTITUTIONS.....	70
METHODS.....	71
“I MAINLY VOLUNTEER BECAUSE IT’S FUN.”.....	73
DIGITAL VOLUNTEERS AND THE SENSE OF COMMUNITY.....	79
“THAT’S A KEY WORD, <i>ENGAGE</i> . THAT’S WHY WE EXIST.”.....	86
CONCLUSION.....	95
BIBLIOGRAPHY.....	97
APPENDICES.....	107
APPENDIX A: “CROWDSOURCING: WHO VOLUNTEERS, AND WHY?”.....	108
APPENDIX B: ADDITIONAL SURVEY RESULTS.....	111
APPENDIX C: RECOMMENDATIONS FOR THOSE CONSIDERING CROWDSOURCING.....	113
APPENDIX D: IRB APPROVAL LETTER FOR SURVEY.....	120
APPENDIX E: IRB APPROVAL LETTER FOR INTERVIEWS.....	121
APPENDIX F: INFORMED CONSENT FOR INTERVIEWS.....	122
APPENDIX G: COPYRIGHT PERMISSIONS.....	124

LIST OF FIGURES

	Page
FIGURE 1: GALAXYZOO IDENTIFICATION PLATFORM.....	52
FIGURE 2: TRANSCRIBE BENTHAM TRANSCRIPTION DESK.....	55
FIGURE 3: DIY HISTORY TRANSCRIPTION PLATFORM.....	56
FIGURE 4: EMPLOYMENT STATUS OF SURVEY RESPONDENTS.....	76
FIGURE 5: LENGTH OF TIME SPENT VOLUNTEERING PER MONTH.....	76
FIGURE 6: MOTIVATIONS FOR VOLUNTEERING.....	78
FIGURE 7: AGE OF SURVEY RESPONDENTS.....	111
FIGURE 8: GENDER OF SURVEY RESPONDENTS.....	111
FIGURE 9: ETHNICITY OF SURVEY RESPONDENTS.....	112
FIGURE 10: ANNUAL HOUSEHOLD INCOME OF SURVEY RESPONDENTS.....	112

CHAPTER I: INTRODUCTION

In his landmark article “The Rise of Crowdsourcing,” journalist Jeff Howe described a new phenomenon he saw taking place on the Internet. Companies like Procter and Gamble, iStockphoto, and others were outsourcing their work to crowds of volunteers or low-paid workers—“crowdsourcing,” Howe dubbed it. He argued that this practice was born out of the open source movement, in which volunteers would collaborate to produce public domain websites and software.¹

More recently, libraries, archives, and museums have also chosen to experiment with the crowdsourcing model. Institutions have found a variety of creative ways to utilize the creativity of the crowd, but some of the most common methods include asking the public to transcribe scans of letters and diaries and to provide metadata and “tags” for images and video. Many institutions have achieved excellent results, both in terms of the data received and in terms of the increased outreach potential with digital volunteers, but other professionals are more skeptical. They wonder if hordes of volunteers on the Internet can really be entrusted with the historical record, and if they could lose their job to these willingly unpaid workers. What even motivates these volunteers to do the work? I argue that most volunteers, both traditional library and archival volunteers and crowdsourcing volunteers, are motivated by a combination of enjoyment of the tasks they complete and a sense of community with other volunteers. Far from the stereotype of the

¹ Jeff Howe, “The Rise of Crowdsourcing,” *Wired* 14 No. 6 (June 2006): accessed October 19, 2013, <http://www.wired.com.ezproxy.mtsu.edu/wired/archive/14.06/crowds.html>.

Internet being a place of lonely strangers, many surveyed crowdsourcing volunteers report feeling a strong sense of community and having forged friendships within the groups they volunteer with.

LITERATURE REVIEW

To understand the nature of crowdsourcing and volunteering on digital platforms, one must first examine literature on more traditional forms of volunteering (both generally, for example within the community, and volunteering in libraries and archives.) It is also important to examine works on the nature of community, as many volunteers on crowdsourcing platforms and forums readily identify themselves as being members of a community. The literature on crowdsourcing in archives and other academic institutions is steadily growing; however most of this writing still consists simply of case studies. Most scholarly writing on crowdsourcing discusses how the process is being used by businesses. Although these works discuss some topics of interest for those in the non-profit sector, other frequently debated subjects, such as the ethics of using free labor for financial gain, are less useful for archivists.

The classic source that most researchers consult first when discussing community associations and organizations is *Democracy in America* by Alexis de Tocqueville. Tocqueville, a French aristocrat, wrote the work in the 1830s after touring the United States. As the title suggests, the book attempts to explain the American democratic system and Tocqueville's interpretation of how and why it works. Tocqueville devotes a part of his work to civil associations, explaining why they are more prevalent in the

United States than in other parts of the world. His interpretation is that in an aristocracy most men have no power but a small number of men have a great deal of power. Thus these men are accustomed to working individually or in very small groups to accomplish their goals. In a democracy, however, no one has much power but everyone has a little power, so everyone must join together to achieve a common goal.²

A more modern publication, *America's Voluntary Spirit*, offers a variety of short essays drawn from everything from the writings of Jane Addams and Andrew Carnegie to scholarly articles written by modern journalists and academics. The authors of the essays have numerous and sometimes conflicting opinions on various topics relating to voluntarism and civil association, but one recurring theme agreed upon by several authors is an explanation for why Americans volunteer and feel such a strong sense of community. Many authors agree that it has to do with Americans' increased physical and social mobility. This explanation is contrary to Tocqueville's theory. In his introduction to the work, editor Brian O'Connell writes, "To portray our history of volunteering as relating solely to goodness may describe the best of our forebears, but it ignores the widespread tradition of organized neighborliness that hardship dictated and goodness tempered."³ O'Connell argues that the first settlers in America had community and

² Alexis de Tocqueville, *Democracy in America*, trans. Arthur Goldhammer, Library of America 147 (New York: Penguin Putnam, 2004), 596.

³ Brian O'Connell, introduction to *America's Voluntary Spirit: A Book of Readings*, edited by Brian O'Connell (New York: Foundation Center, 1983), xix.

family structures that differed from other countries and that forced them to be interdependent.⁴

Authors of other essays within the book agree. In his essay “The Joiners” (an excerpt from the book *America as a Civilization*), journalist Max Lerner argues that American’s ability to navigate the social ladder creates a greater need for community. He notes that older, hierarchical societies have less need for associations because everyone “knows their place” and stays there. But in the United States, one can move up and down the social ladder and is not defined only by his or her social class. Therefore, Americans must make connections to define themselves by their interests and personalities. Historian Daniel J. Boorstin makes a similar argument in an essay excerpted from his book *The Decline of Radicalism*. He writes that in nineteenth century Europe, most people lived where they were born and where their family had been living for centuries. In the United States, however, everyone (or one of their recent ancestors) had recently made a decision about where to live. “The sense of community,” he writes, “was inevitably more vivid and more personal because, for so many in the community, living here had been an act of choice.”⁵

⁴ Ibid., xix-xx.

⁵ Max Lerner, “The Joiners,” in *America’s Voluntary Spirit: A Book of Readings*, edited by Brian O’Connell (New York: Foundation Center, 1983), 82; Daniel J. Boorstin, “From Charity to Philanthropy,” in *America’s Voluntary Spirit: A Book of Readings*, edited by Brian O’Connell (New York: Foundation Center, 1983), 131.

Another immensely important work in the literature on association and community engagement is *Bowling Alone: The Collapse and Revival of American Community* by Robert D. Putnam. Writing in 2000, Putnam argues that since his childhood America has been seeing less and less community engagement of all forms, including political engagement (voting, running for office), volunteer activities, and even social activities such as bridge clubs and bowling leagues. He further explains that this disengagement has negative effects on individuals and on society as a whole, and ends the book with ideas for possible “cures” for the lack of civic engagement.⁶ Putnam believes that several factors coincide to explain the reduction in civic engagement, but the factor that he states to be the most important is that of generational change. He argues that the generation of Americans who grew up during the Great Depression and fought in World War II, whom he calls the “long civic generation,” learned the importance of community and civic engagement from these disastrous events. Later generations who grew up in relative comfort have never had such an eye-opening example of the need for community association.⁷

Little has been written on the history of voluntarism, especially about the history of volunteering in libraries and archives. Most work on this subject focuses on the Progressive Era and examines the struggles of the progressives to change the world around them. Many authors pay particular attention to women’s roles in progressive era

⁶ 1. Robert D. Putnam, *Bowling Alone: The Collapse and Revival of American Community* (New York: Simon & Schuster, 2000), 27-28.

⁷ *Ibid.*, 283-284.

reform and volunteer groups as a way to describe the power struggles brought on by gender. Just a few examples of this thread of historiography are *African American Women and Social Action: The Clubwomen and Volunteerism from Jim Crow to the New Deal, 1896-1936* by Floris Loretta Barnett Cash, and *Relations of Rescue: The Search for Female Moral Authority in the American West, 1874-1939* by Peggy Pascoe.⁸ These works argue that women were involved in voluntarism and in great numbers during the progressive era because it was a way of having power outside of the home. Women of the time often found themselves educated beyond what was typical and expected for their gender and race and in seeking a method to utilize their talents, turned to voluntarism. Although they felt a sense of responsibility towards the poor whom they were helping, women also hoped to achieve some amount of power through their volunteer organizations, though how much power varied.

Very few works discuss the history of volunteering in libraries, and those that do often mention it offhand; instead most authors choose to discuss case studies and recommendations for the use of volunteers. For example, in his paper concerning the beginnings of the archival profession, Waldo Leland noted that many early “descriptive accounts” of records, probably finding aids, were written by volunteer historians.⁹

⁸ Flora Loretta Barnett Cash, *African American Women and Social Action: The Clubwomen and Volunteerism from Jim Crow to the New Deal, 1896-1936* (Westport: Greenwood Press, 2001); Peggy Pascoe, *Relations of Rescue: The Search for Female Moral Authority in the American West, 1874-1939* (New York: Oxford University Press, 1990).

⁹ Waldo Gifford Leland, “The First Conference of Archivists, December 1909: The Beginning of a Profession,” in *American Archivist* 13 no 2 (1950): 111, accessed April 7, 2014, doi: 10.17723/aarc.13.2.h874j87h80441422.

Another article, “Volunteers in Libraries: Program Structure, Evaluation, and Theoretical Analysis” gives only a brief overview of the history of volunteers in libraries, instead spending most of the article describing how best to utilize library volunteers.¹⁰

Archival literature is even sparser in its discussion of volunteers. Most archival literature discussing volunteers consists of case studies, such as two articles from the National Archives and Records Administration magazine entitled “NARA’s Armies of Volunteers” and “Our Wonderful Volunteers,” which discuss the vast numbers of volunteers the institution utilizes, the types of projects they work on, and the training they undergo.¹¹ These two articles offered advice for other archives implementing volunteer programs, but the primary goal seemed to be to congratulate current volunteers. An older but more detailed case study titled “Using Volunteers for Special-Project Staffing at the National Air and Space Museum Archives” discusses a project undertaken by the National Air and Space Museum (NASM) in which the museum recruited volunteers from across the country to attend one of several two-week sessions volunteering in the

¹⁰ Erica A. Nicol and Corey M. Johnson, “Volunteers in Libraries: Program Structure, Evaluation, and Theoretical Analysis,” in *Reference & User Services Quarterly*, 48 no 2 (2008): 154-155, accessed April 7, 2014, <https://ezproxy.mtsu.edu:3443/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=aph&AN=35665049&site=eds-live&scope=site>.

¹¹ Lee Ann Potter and Rebecca Martin, “NARA’s Armies of Volunteers,” in *Prologue* 38 no 4 (Winter 2006), accessed January 13, 2016, <https://www.archives.gov/publications/prologue/2006/winter/volunteers.html>; Adrienne C. Thomas, “Our Wonderful Volunteers,” in *Prologue* 41 no 3 (Fall 2009), accessed January 12, 2016, <https://www.archives.gov/publications/prologue/2009/fall/archivist.html>.

museum's archives.¹² This article's stated goal was to offer advice to archives considering a similar short-term volunteer program. Although these articles are interesting, they are of little use to small libraries and archives that do not have the resources for the intensive training sessions of NARA or the prestige necessary to recruit volunteers across the country for short term volunteering sessions. More useful is "Best Practices for Volunteers in Archives" published in 2014 by the Society of American Archivists, which is simply a short guide meant to provide advice and further resources for archivists and volunteers at institutions with volunteer programs.¹³ This thesis will fill a gap by offering advice for managing crowdsourcing projects and volunteers that is currently lacking in archival literature.

Unlike voluntarism, there has been a large amount of material written on crowdsourcing. However, it is found spread across the reading of various disciplines. Most references, unsurprisingly, are found in technology journals, but others are found in business and library science journals or simply scattered across the blog posts of journalists. Most authors of articles on crowdsourcing argue that it is a new idea that came about because of the Internet and its collaborative nature. Journalist Jeff Howe asserts this in his article "The Rise of Crowdsourcing," which first named the

¹² Susan E. Ewing, "Using Volunteers for Special-Project Staffing at the National Air and Space Museum Archives," in *American Archivist* 54 no 2 (1991): 176-183, accessed January 12, 2016, <http://www.jstor.org/stable/40293550>.

¹³ Society of American Archivists, "Best Practices for Volunteers in Archives." *Society of American Archivists*, August 2014, accessed January 14, 2016, <http://www2.archivists.org/standards/best-practices-for-volunteers-in-archives>.

phenomenon, and other authors have agreed with him, notably Thomas Goetz in his article “Open Source Everywhere.”¹⁴

A major point of contention between those writing on crowdsourcing is whether the activity is helpful, or whether it creates too much work for those soliciting the work (such as archivists) with unsatisfactory results. Most archivists agree that it is useful, as is argued in articles like “For Bentham and Others, Scholars Enlist Public to Transcribe Papers” by Patricia Cohen, “The Rise of Crowdsourcing” by Jeff Howe, and “Crowdsourcing: How and Why Should Libraries Do It?” by Rose Holley.¹⁵ Some scholars, however, argue that crowdsourced projects are not to be trusted because the volunteers do not have the necessary expertise to complete them and therefore provide inaccurate results. Examples include technology blogger Nick Douglas, who compares crowdsourcing to serfdom, Péter Jascó, who writes that Wikipedia looked like “a joke at best,” and Andrew Keen, whose book *The Cult of the Amateur: How Blogs, MySpace, YouTube, and the Rest of Today’s User-Generated Media Are Destroying Our Economy*,

¹⁴ Howe, “Rise of Crowdsourcing”; Thomas Goetz, “Open Source Everywhere,” *Wired* 11, no. 11 (November 2003), accessed October 20, 2013, http://www.wired.com.ezproxy.mtsu.edu/wired/archive/11.11/opensource_pr.html.

¹⁵ Patricia Cohen, “For Bentham and Others, Scholars Enlist Public to Transcribe Papers,” *New York Times* December 27, 2010, accessed October 20, 2013, <http://www.nytimes.com/2010/12/28/books/28transcribe.html>; Howe, “Rise of Crowdsourcing,” Rose Holley, “Crowdsourcing: How and Why Should Libraries Do It?” in *D-Lib Magazine*, 16, no 3/4 (March/April 2010).

Our Culture, and Our Values argues that not only crowdsourcing but all Web 2.0 media is destroying today's culture.¹⁶

This investigation will build on all these threads by studying the various definitions of voluntarism put forth by different authors and studying the motivations of volunteers. I will also research the history of crowdsourcing in archives in order to compare the nature of the crowdsourcing community to the community ties that result from volunteering in a more traditional sense. I will also provide a more thorough answer to the debate surrounding the usefulness of crowdsourcing and provide recommendation for institutions considering adopting crowdsourcing projects.

Crowdsourcing in archives is important to study because archives are underfunded and understaffed. Crowdsourcing allows for the completion of the projects that would otherwise be too expensive. These projects, in turn, will allow greater public access to archival resources. For instance, transcription completed by crowdsourcing can allow for full text searches of documents, making currently inadequately described collections more accessible. Likewise, crowdsourcing the work of adding metadata can also make digital collections more accessible by allowing the public to search using familiar terms, rather than struggling with unfamiliar professional archival language.

¹⁶ Nick Douglas, "Job Market News: That's Not Slave Labor, That's Crowdsourcing!" in *Valleywag [Gawker] Media blog* May 25, 2006 4:46, accessed October 19, 2013; Péter Jascó, "Péter's Picks & Pans," *Online Magazine* 26, no. 2 (April 2002): 79–82, accessed November 1, 2013, <https://ezproxy.mtsu.edu:3443/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=llf&AN=502875163&site=eds-live&scope=site>; Andrew Keen, *The Cult of the Amateur: How Blogs, MySpace, YouTube, and the Rest of Today's User-Generated Media Are Destroying Our Economy, Our Culture, and Our Values* (New York: Random House, 2008).

Finally, crowdsourcing projects benefit the volunteers by allowing individuals from around the world to interact with archival records rather than simply passively viewing those records.

METHODOLOGY

This thesis will cover the late nineteenth century to the present in its mission to tie crowdsourcing to historical ideas of volunteering. To do so, I examine a mix of historical, sociological, and technological sources to understand why people chose to volunteer historically, what the history of crowdsourcing is, and how crowdsourcing connects historically to voluntarism. Because crowdsourcing is a current phenomenon and is still being written about, I used many sources as both primary and secondary sources. Those used as primary sources include articles from archives, technology, law, and library journals, crowdsourcing websites, and technology blogs.

I also conducted surveys and interviews as primary sources. I designed and received IRB approval for both a survey and for interviews. The survey was to determine the demographic information and the motivation for volunteering among those who volunteer on academic crowdsourcing websites. The surveys were sent to those who volunteer on the crowdsourcing websites DIY History and Project Gutenberg Distributed Proofreaders, and asked demographic questions and asked volunteers how much time they spent volunteering and why they chose to volunteer their time on crowdsourcing websites. I use the results from these surveys to compare the motivation of modern crowdsourcing volunteers to traditional volunteers. In addition, I conducted the

interviews with staffers who worked with anonymous crowdsourcing websites, asking them how successful quantitatively their project had been, whether they believed their project had been a success, and whether they viewed crowdsourcing as worthwhile or unsatisfactory and whether they would recommend it to other archivists. I use these interviews, along with scholarly articles on the merit of crowdsourcing, to argue crowdsourcing's usefulness.

CHAPTER OUTLINE

Chapter one examines the history of traditional volunteering in America (volunteering that takes place in person, as opposed to on the Internet), in libraries and archives as well as in other volunteer organizations. In this chapter I focus on the motivations of volunteers and those who recruit them, as well as how scholars from various disciplines, such as economics, sociology, or political science, explain these motivations. This information will later be examined alongside similar information from participants in crowdsourcing projects to compare their similarities and differences.

In chapter two, I recount the history of crowdsourcing—both crowdsourcing and Internet voluntarism in general and the more narrow history of crowdsourcing in libraries, archives, and other academic fields—as well as the histories of certain key projects. As many authors have different definitions of crowdsourcing, this chapter provides a working definition for use for the remainder of the thesis. It closes by discussing the criticisms that many researchers have leveled against crowdsourcing and Web 2.0.

Chapter three examines the results of both the survey sent to crowdsourcing volunteers, as well as interviews with staff of crowdsourcing projects in order to show the connections between traditional volunteering and crowdsourcing, as well as to show the successes that can come from crowdsourcing. I argue that traditional archival volunteers and crowdsourcing volunteers are both motivated by a combination of enjoyment of the tasks they perform and a strong sense of community and friendship shared among volunteers.

CHAPTER II:

MOTIVATIONS OF VOLUNTEERS

Why would an individual willingly choose to give away hours of free time that could be spent on leisurely pursuits? This question, along with the simple question of how to define the word “volunteer,” has been difficult for academics to answer. Does any form of unpaid labor count as volunteering, or must other conditions be met? Do individuals volunteer altruistically, or do all volunteers receive some sort of benefit, whether tangible or intangible, in return? This chapter lays groundwork for the study of digital volunteering by examining how various authors have sought to answer these questions and by investigating the motivations and the types of tasks performed by those who volunteer in person in archives. Scholars of different disciplines differ profoundly in how they define volunteering and about what motivates volunteers. The motivations of historical and archival volunteers align most closely with those descriptions from authors who define volunteering as a leisurely and pleasurable activity that is motivated by a combination of self-interest and altruism.

DEFINING “VOLUNTEERING”

The question of how volunteers and volunteer work should be defined is a primary question that researchers must answer. Works by historians often are not very helpful to study when looking for definitions because rather than defining volunteering as a theory, historians often focus on individual voluntary groups. Because the discipline studies society and social interactions, sociological texts are extremely useful for offering

theories and definitions, however literature from many other disciplines also offers valuable insights.

“Navigating Theories of Volunteering: A Hybrid Map for a Complex Phenomenon” explores and attempts to synthesize several different theories of volunteering. In this work, sociologists Lesley Hustinx, Ram A. Cnaan, and Femida Handy cite previous work by Cnaan, Handy, and M. Wadsworth, who, in turn, had examined two hundred definitions offered by other sociologists and found that all definitions included some mention of “time, labor, [and] expertise.” Hustinx, Cnaan, and Handy note, however, that rather than describing what volunteering is, most definitions describe what volunteering is not: paid, forced, etc. In their work, the authors therefore examine how volunteering has been studied and explained by others across disciplines.¹

Economists simply define volunteering as “unpaid work” with a value that can be mathematically determined, and argue that volunteers’ main motivations are to acquire usable skills. To economists, “volunteering is a paradox...they [volunteers] undertake activities wherein their costs exceed their benefits.”² As such, economists have developed several models to explain the benefits individuals receive from volunteering in an attempt to explain the paradox. According to these models, volunteers may receive private benefits, skills, public goods and services, or simply the good feeling that volunteering

¹ Lesley Hustinx, Ram A. Cnaan, and Femida Handy, “Navigating Theories of Volunteering: A Hybrid Map for a Complex Phenomenon,” *Journal for the Theory of Social Behaviour* 40 no. 4 (December 2010): 412-415, accessed April 23, 2014, doi: 10.1111/j.1468-5914.2010.00439.x.

² Ibid., 411, 415.

brings. Volunteers, therefore, are “impure altruists...interested in both private and public benefits of volunteering.”³ Sociologists and political scientists, however, view volunteering as an integral part of society, democracy, support, and harmony. Political scientists in particular believe volunteering to be a civic duty. Sociologists do not shy away from self-interest, noting that for some volunteers, it can be a stepping stone on the way to a career (such as an unpaid internship), however according to Hustinx, Cnaan, and Handy, sociologists embrace a more altruistic view of volunteering. They write:

It is considered an essential and exceptional form of social solidarity that binds society together. The act of volunteering stands out as a primary expression of core human values such as altruism, compassion, concern for others, generosity, social responsibility, and community spirit...It is a fundamental expression of community bonding and group identity.⁴

Psychologists relate volunteering to their study of personality and attempt to discover what personalities tend to volunteer, and what personality traits volunteers have in common. Some of these personality traits are “social value orientation, empathic concern, perspective taking, self-efficacy, and positive self-esteem.”⁵

Sociologist Robert A. Stebbins presents a completely different definition of volunteering. He argues that “the reigning conception—volunteering as unpaid labor” was nowhere close to what was actually taking place, and that instead volunteering

³ Hustinx, Cnaan, and Handy, “Navigating Theories of Volunteering,” 415-416.

⁴ Ibid., 411, 419, 418, 417.

⁵ Ibid., 418-419.

needed to be studied as a form of leisure.⁶ Stebbins' definition of volunteering is "uncoerced help offered either formally or informally with no, or at most, token pay, done for the benefit of both other people and the volunteer."⁷ He notes that in order for this definition to work, volunteering cannot be coerced, but he also places caveats on what constitutes true coercion and obligation. For example, if an individual is an officer in a club, then they are required to attend meetings, but if that individual signed up because they enjoy the activity then the obligation to attend is less important than the enjoyment of attendance. As a support to his argument, Stebbins notes that many surveys of volunteers have indicated that they consider their volunteer work to be a form of leisure.⁸ Because volunteering is a leisure activity, self-interest must play a role. Altruism is not the sole motivation for volunteering. Within this definition, Stebbins states that there are three different types of volunteering: serious or career, casual, and project. Career volunteering involves volunteering over a long period of time and utilizing special abilities that the individual may possess, such as coaching a Little League baseball team. Casual volunteering includes simple acts that do not require much training and are "immediately, intrinsically rewarding" such as "cooking hot dogs at a church picnic." Project volunteering is a one-time event that requires a large amount of skill or planning, but is not meant to turn into a long term volunteering opportunity, for

⁶ Hustinx, Cnaan, and Handy, "Navigating Theories of Volunteering," xiii.

⁷ Robert A. Stebbins, introduction to *Volunteering as Leisure/Leisure as Volunteering: An International Assessment*, ed. Robert A. Stebbins and Margaret Graham, (Cambridge: CABI Publishing, 2004), 5.

⁸ *Ibid.*, 4, 7-10.

example planning a party.⁹ Stebbins offers a definition that is very different from other scholars of volunteerism, yet also similar. Describing volunteering as leisure is a new perspective, however this definition allows the authors to embrace both the motivations of self-interest and altruism that many other authors struggle to separate.

MOTIVATION BEHIND FEDERAL VOLUNTEER PROGRAMS

Stebbins writes that “volunteering is, among other things, a primarily creative, society-building activity, which nevertheless loses this quality when, as a money-saving strategy, it is foisted on altruistic citizens by agents of the public or private sector.”¹⁰

Despite this condemnation, those enlisting the help of volunteers have a variety of reasons for doing so, which can in turn be either altruistic or selfish.

One extremely large-scale volunteer project was the army’s decision to stop drafting recruits and switch to an “all-volunteer force,” consisting of recruits who had chosen long-term employment in the military rather than fulfilling their wartime duty. When the army decided to eliminate the draft and switch to an “all-volunteer force,” their motivations were altruistic because they benefitted average Americans and yet self-interested because they led Americans to view the military in a more favorable light. Beth Bailey writes about some of the army’s recruitment methods to try to convince soldiers to enlist of their own accord and create an “all-volunteer force,” rather than reinstating the draft. Although joining the military could be seen as a civic duty, the draft infringed on

⁹ Stebbins, introduction, 5-6.

¹⁰ Ibid., 2-3.

the individual liberties Americans held so dear. The primary method the army used was economic incentive and making the army an attractive alternative career choice for those graduating from high school. One recruiter created the “Be All That You Can Be” recruitment slogan that turned out to be a boon to the army. Commercials and advertisements under this slogan were designed to show that recruits would gain skills and benefits that they could later use in a civilian job or to attend college. This slogan was a great recruitment tool because it promoted positive feelings toward the army and encouraged recruits to fulfill their potential.¹¹ This recruitment strategy played into a recruit’s self-interest: volunteering for the armed forces was no longer about civic duty but was instead simply an alternative career choice, and one that was increasingly attractive in its benefits.

The founding of another federally run volunteer program, the Peace Corps, created a debate over whether it was an idealistic volunteer group or a ploy of United States foreign policy. Elizabeth Cobbs Hoffman explains the Peace Corps “sought a meeting point for both the crudest and the finest national interests, from military security and the creation of wealth to fulfillment of the philosophical ideas of the Declaration of Independence.”¹² One question Hoffman raises is whether government could adequately use humanitarianism for honest purposes, or whether it was actually used to help expand the nation’s hegemony. The United States government was in constant struggle over how

¹¹ Beth Bailey, *America’s Army: Making the All-Volunteer Force* (Cambridge: Belknap Press, 2009), 194-196.

¹² Elizabeth Cobbs Hoffman, *All You Need Is Love: The Peace Corps and the Spirit of the 1960s* (Cambridge: Harvard University Press, 1998), 7.

the Peace Corps should be used. One example is the state department's desire to send Peace Corps volunteers to contested areas, such as Vietnam. Peace Corps leaders refused, because Dean Rusk as argued, "The Peace Corps is not an instrument *of* foreign policy because to make it so would rob it of its contribution *to* foreign policy."¹³ Sargent Shriver, director of the Peace Corps, felt that "the Peace Corps was not an 'arm' or 'tool' of the Cold War, [but] nevertheless deeply believed that the Peace Corps would help the United States win it."¹⁴

Both of the above examples illustrate how the motivations of those recruiting volunteers can vary, and how one organization can have multiple reasons for recruiting volunteers. The leaders of the army and the Peace Corps had both well-intentioned and problematic reasons for recruiting volunteers. For the army, preserving individual freedom by eliminating the draft was the result of increasing opposition to conscription. Taking this step to end conscription was altruistic in the sense that it was beneficial to those citizens who wanted to avoid military service and those volunteers who benefited from increased pay. The military itself, however, also benefited from the end of conscription because it was viewed more favorably in the public eye and attracted more recruits. The Peace Corps' main motivation was humanitarianism and education for nations around the world, but as Hoffman points out, the United States government was also motivated by hopes that the humanitarian work of volunteers around the world would lead to other countries viewing the United States in a more favorable light during

¹³ Hoffman, *All You Need Is Love*, 4, 99.

¹⁴ *Ibid.*, 106.

the Cold War. Although public sector organizations are less likely to take advantage of volunteers than private sector organizations, who make a profit off of the free labor, these examples show that public sector organizations are not immune, and should be cautious of their motivations.

MOTIVATIONS OF VOLUNTEERS

There are many reasons that an individual could choose to volunteer for a project, but many people, scholars and otherwise, do not understand how this could be the case. Hustinx, Cnaan, and Handy write that there is a “problem of collective action”—no one wonders why people want a job, but they wonder why people volunteer. Volunteering is considered irrational and more researchers study motivation to volunteer than how much people volunteer or other aspects of volunteering.¹⁵ Motivation is important to study because as mentioned earlier, volunteering is somewhat of a paradox. Although some scholarly disciplines try to remove self-interest as a potential motivation, it is in fact quite difficult to separate altruism and self-interest because even if an individual volunteers because it gives them a stereotypical warm fuzzy feeling, this is, in fact, receiving a benefit. Therefore most volunteers, both modern and historic, are motivated by a combination of altruism and self-interest.

According to theologian Anne Birgitta Yeung, motivation is vital to the study of volunteering—“individual motivation is the core of the actualization and continuity of

¹⁵ Hustinx, Cnaan, and Handy, “Navigating Theories of Volunteering,” 420.

volunteer work.”¹⁶ Motivation is what causes individuals to volunteer and continue to do so. As a result of her conversations with volunteers, Yeung found “forty seven motivational themes” that fell under eight (sometimes contradictory) categories: getting, giving, action, thought, proximity, distance, newness, and continuity. She plotted these eight themes on a chart shaped like a star so that the themes can be interconnected, and it is even possible to create a sort of map out of a single volunteer’s various reasons for choosing to volunteer. Yeung hopes that this “octagon model might shed light on theoretical issues such as the dilemma of the multilayered nature of altruism.”¹⁷ Although Yeung claims that she does not seek to explain volunteering, her article does just that with its model. Like several other authors, she blends the ideas of altruism and self-interest (the “getting” section of her model), however she makes it clear that she does not feel that self-interest is a drawback of altruism.

Sociologists argue several theories on what motivates individuals to volunteer. Stebbins argues that motivation is not especially important when defining volunteering as unpaid labor, but that it is especially important when defining it as leisure. Different demographic groups have different motivations for volunteering. However, similar to other scholars, he writes all groups are motivated by “altruism and self-interest.”¹⁸ Self-

¹⁶ Anne Birgitta Yeung, “The Octagon Model of Volunteer Motivation: Results of a Phenomenological Analysis,” *Voluntas: International Journal of Voluntary & Nonprofit Organizations* 15, no. 1 (March 2004): 21, accessed February 29, 2016, <https://ezproxy.mtsu.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=aph&AN=12817951&site=eds-live&scope=site>.

¹⁷ Ibid., 23, 32-33, 40, 41.

¹⁸ Stebbins, introduction, 2.

interest can include volunteering for some sort of reward or can simply be volunteering for a cause one feels strongly about. He also identifies three different types of volunteers, based on their motivation. “Active motivators” want to give back to the community or to be challenged. “Passive motivators” are people who were attracted to a volunteering opportunity by social media or a friend or family member. This group usually enjoys volunteering because they find it enjoyable, it teaches them new skills, and they can meet new people. Finally, “special interest motivators,” are attracted to a volunteer opportunity simply because they are interested in the project at hand, such as a “history buff” volunteering to lead museum tours.¹⁹

Two examples from the expansive history of progressivism and settlement houses offer interesting perspectives on volunteer motivation. African American women who founded settlement houses (institutions that provided shelter, education, and cultural opportunities to tenants and the surrounding neighborhood) at the turn of the twentieth century were also motivated by a combination of altruism and self-interest. African American settlement house workers volunteered because they were “educated beyond their race and gender” (compared to other women of the time and to the lives they were expected to lead) and “had a sense of responsibility.”²⁰ Floris Barnett Cash argues that the time period did not allow married women and mothers to pursue employment, and so voluntary work became a way for women to use their skills for a use other than the home

¹⁹ Ibid., 3-4, 25-26.

²⁰ Floris Barnett Cash, *African American Women and Social Action: The Clubwomen and Volunteerism from Jim Crow to the New Deal, 1896-1936* (Westport: Greenwood Press, 2001), 3.

life. Cash also states that they tried to advance their race by opposing segregation, lynching, and other forms of racism.²¹ Volunteering for these women was primarily a question of self-interest and wanting to make use of their education, however this self-interest led them to pursue altruistic work.

Progressive Era religious women in the western United States were similar in their self and public interest motivation in their building of rescue homes for other women. They were guided primarily by self-interest and their own wishes; however by changing society in their favor they could help all women. They feared that the sins of men, especially those sexual in nature, could be destructive, and so they sought to create matriarchal Christian homes in the form of rescue homes to criticize men's power. They were appalled by men's power over western cities and still held Victorian beliefs that women were virtuous, so they chose to rescue women from the abuse of men. They hoped to gain moral authority and influence over men in the process, however most of their power ended up being over the women who came to reside in the rescue homes.²² Similar to African American founders of settlement houses, the self-interest of religious women spurred them to altruistic action in creating religious rescue homes.

The Peace Corps was slightly different in that there were both altruistic and self-interested motivations involved. These mixed feelings were not necessarily all felt by volunteers, however. Hoffman argues that "The United States was the first nation...to

²¹Cash, *African American Women and Social Action*, 9, 5.

²² Peggy Pascoe, *Relations of Rescue: The Search for Female Moral Authority in the American West, 1874-1939* (Oxford: Oxford University Press, 1990), 36-37, xvi, 33-34.

incorporate volunteering into its foreign policy in an attempt to demonstrate one alternative to power politics.”²³ The second Peace Corps director said that “the Peace Corps is about love,” but in some aspects it was more about assuaging the feelings of Americans.²⁴ The Peace Corps was founded during upheaval to reassure Americans that America was, in fact, good and was a “superhero, protector of the disenfranchised, [and] defender of the democratic faith.”²⁵ It was meant to spread the best parts of American ideals and character around the world, and to form an international community based on “universal values.”²⁶ Hoffman concluded that “perhaps training its own citizens as better servants of humanity was itself enough reason to send Americans abroad,”²⁷ noting that volunteers rated their experiences as worthwhile and most went on to have similar “helping” careers such as teachers, congressmen, and ambassadors, and to work in the nonprofit sector.²⁸ The Peace Corps, then, was a mixture of self-interest and altruism—altruism on the part of the volunteers who were sent around the world and self-interest from the government, who recognized the impact such a program could have on the global community and hoped that the good will the Peace Corps spread would shine favorably back upon the United States.

²³ Hoffman, *All You Need Is Love*, 8-9.

²⁴ *Ibid.*, 9.

²⁵ *Ibid.*, 1.

²⁶ *Ibid.*, 23-24.

²⁷ *Ibid.*, 250.

²⁸ *Ibid.*, 257.

All of the examples above describe motivations that are a combination of altruism and self-interest. As Yeung and Stebbins point out, although altruistic acts are normally defined to be entirely selfless, self-interest and altruism are frequently difficult to separate. Women building settlement houses may have been primarily motivated by the opportunity to gain social power, but they also felt a responsibility to use their education to help others of their race. Peace Corps members, on the other hand, volunteered primarily out of altruism, but gained valuable experience that led many of them to successful careers in the public sector. As these examples show, it seems difficult for any volunteer to have absolutely no interest in their work and to volunteer their time purely out of selflessness.

VOLUNTEERS IN LIBRARIES AND ARCHIVES

The American Library Association (ALA) and the Society of American Archivists (SAA) have both adopted recommendations for the use of volunteers, although both have varying advice on how to handle volunteers. The SAA's guidelines, known as the "best practices," were adopted in June of 2014, while the ALA's guidelines have not been updated since 1971.²⁹ The biggest aspect that both organizations agree on is that volunteers should not replace paid staff.³⁰ The SAA also notes that institutions should

²⁹ Society of American Archivists, "Best Practices for Volunteers in Archives," *Society of American Archivists*, August 2014, 1, accessed January 14, 2016, <http://www2.archivists.org/standards/best-practices-for-volunteers-in-archives>; "ALA Standards & Guidelines," American Library Association, March 26, 2015, accessed February 19, 2016, <http://www.ala.org/tools/guidelines/standardsguidelines>.

³⁰ Society of American Archivists, "Best Practices for Volunteers in Archives," 2; "Guidelines for Using Volunteers in Libraries," *American Libraries* 2, no. 4 (1971): 407–8, accessed February 19, 2016, <http://www.jstor.org/stable/25618274>.

consider adopting written policies and procedures to apply to volunteers and that private and for-profit organizations should refrain from utilizing volunteers.³¹ The ALA takes these recommendations further to say that rather than considering written policies, libraries *must* adopt them, as well as offering training and detailed job descriptions to potential volunteers. Volunteers should also be given work that they feel comfortable with and that utilizes their strengths and interests.³²

Those who volunteer in libraries and archives also have a variety of motivations for their volunteer work. But, many fall into one of two categories described by Stebbins: “passive motivators,” who find the activity enjoyable and enjoy the social aspect and the people they meet, or “special interest motivators,” who, like history buffs, are mostly interested in the project at hand. Volunteers at the National Archives and Records Administration (NARA) note that they love “the social part” of volunteering at the National Archives, and that their friendships with staff and other volunteers motivate them to keep coming back.³³ Other NARA volunteers became involved specifically out of an interest in the records, such as genealogists who volunteer to help others with their genealogical research, or military veterans who volunteer specifically to process old

³¹ Society of American Archivists, “Best Practices for Volunteers in Archives,” 2-3.

³² “Guidelines for Using Volunteers in Libraries,” 407-408.

³³ Stebbins, introduction, 3-4, 25-26; Lee Ann Potter and Rebecca Martin, “NARA’s Armies of Volunteers,” in *Prologue* 38 no 4 (Winter 2006), accessed January 13, 2016, <https://www.archives.gov/publications/prologue/2006/winter/volunteers.html>

military records.³⁴ Some university archivists note that because they have a strong interest in the work, “history buff” volunteers are more accurate workers than work-study students from outside of the history field.³⁵

For library volunteers, enjoying their work can be a huge motivation. Librarians Bonnie F. McCune and Cheryl A. McHenry speak to the need of matching volunteers to assignments based on their skills and interests. McHenry also argues that knowing an individual’s motivation for volunteering can help when placing them in a role.³⁶ McCune would likely agree with this argument, as she notes several examples of volunteer placement gone wrong—for example, a “career woman” who wants to promote childhood literacy who is offered a role helping out with weekday story times, or “a poetry-loving retired business executive, once responsible for hundreds of employees,” who is assigned to addressing envelopes.³⁷ According to McHenry, “Motivation depends on successfully satisfying one’s desires or goals. A simple first step is to inquire about the

³⁴ Adrienne C. Thomas, “Our Wonderful Volunteers,” in *Prologue* 41 no 3 (Fall 2009), accessed January 12, 2016, <https://www.archives.gov/publications/prologue/2009/fall/archivist.html>.

³⁵ Rhonda Huber Frevert, “Archives Volunteers: Worth The Effort?,” *Archival Issues* 22, no. 2 (1997): 149, accessed February 28, 2016, <http://www.jstor.org/stable/41101978>.

³⁶ Cheryl A. McHenry, “Library Volunteers: Recruiting, Motivating, Keeping Them,” *School Library Journal* 34, no. 9 (May 1988): 46, accessed February 19, 2016, <https://ezproxy.mtsu.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=aph&AN=5770264&site=eds-live&scope=site>.

³⁷ Bonnie F. McCune, “The New Volunteerism: Making It Pay Off For Your Library,” *American Libraries* 24, no. 9 (1993): 822, accessed February 19, 2016, <http://www.jstor.org/stable/25633036>.

volunteer's likes or dislikes to best understand how to motivate them."³⁸ In this sense, then, it can be argued that a large part of motivating and maintaining volunteers is simply assigning tasks that the individual finds enjoyable.

Volunteers in archives can take on a variety of different tasks, and the level of difficulty of those tasks depends simply on the level of trust and training a given institution is able to afford their volunteers. For example, when the National Air and Space Museum archives invited members of the public for two week sessions of volunteering they had minimal time for intense training, and therefore offered simple projects for their volunteers to complete. These projects included indexing, sorting, labeling, rehousing, writing photograph descriptions, and "using a checklist of possibilities to determine preservation needs."³⁹ NARA, however, provides much more training to their volunteers and entrusts them with a wider variety of tasks. Volunteers are expected to complete at least sixteen hours of training, and those who lead tours are required to complete an additional sixty hours of training. Although NARA volunteers complete some simple tasks, such as indexing, labeling, and rehousing, they are often entrusted with more difficult tasks including reference help, writing research guides and translations, and working as tour guides.⁴⁰ Volunteers can also be great public relations advocates for libraries and archives, both in terms of promoting use and advocating for

³⁸ McHenry, "Library Volunteers," 46.

³⁹ Susan E. Ewing, "Using Volunteers for Special-Project Staffing at the National Air and Space Museum Archives," *American Archivist* 54 no 2 (1991): 182, accessed January 12, 2016, <http://www.jstor.org/stable/40293550>.

⁴⁰ Potter and Martin, "NARA's Armies of Volunteers."

funding and legislation.⁴¹ Professionals advocating for funds for their own institution may be accused of simply being worried about losing their jobs, however volunteers can promote an institution without any such concern.

As of this writing, libraries and archives are currently looking to recruit volunteers for a wide variety of tasks. Libraries often want help working with the public, and archives want help with “behind-the-scenes” tasks. The Des Moines Public Library utilizes volunteers to assist with the summer reading challenge, help non-native English speakers practice their language skills, and assist with the preparation for and the hosting of various library programs.⁴² San Diego County Library volunteers read to adults and children, help children with homework, and host book discussions.⁴³ As of this writing, the Jefferson County Public Library in Colorado is actively searching for volunteers to assist on the bookmobile and host book clubs. The library’s volunteer website notes that the March 2016 Volunteer of the Month was chosen because she took over deliveries on several bookmobile and home delivery routes that had been abandoned by other volunteers.⁴⁴ The Maryland State Archives utilizes volunteers for both public and behind-the-scenes activities. According to the website, their volunteer options fall into two major categories: “Appraisal and Description Volunteers” and “Reference Volunteers.”

⁴¹ Frevert, “Archives Volunteers,” 149-150.

⁴² “Volunteer,” Des Moines Public Library, accessed March 5, 2016, <http://dmpl.org/volunteer>.

⁴³ “Volunteer Opportunities,” San Diego County Library, accessed March 5, 2016, <http://www.sdcl.org/volunteer.html>.

⁴⁴ “Volunteer,” Jefferson County Public Library, 2016, accessed March 5, 2016, <http://jeffcolibrary.org/volunteer>.

Appraisal and Description Volunteers work on tasks such as indexing, labeling folders, removing staples, flattening, data entry, and research. Reference Volunteers commit to regular shifts at the reference desk to provide basic orientation for new researchers such as teaching researchers to fill out request slips and how to read catalogs and finding aids.⁴⁵ The Smithsonian has an established volunteer program called the “Behind-the-Scenes Volunteer Program” allowing individuals to volunteer in the institutions many museums and archives. Library and archives volunteers at the Smithsonian typically catalog, shelve, and organize materials, with library volunteers also compiling annotated bibliographies. Multilingual speakers also have the option of volunteering to translate documents.⁴⁶

Although most volunteers display a mix of self-interest and altruism, many library and archives volunteers can lean further away from the altruism when they are volunteering in the hopes of acquiring skills to list on a resume. Several authors list this as the most common motivator for library and archives volunteers, and it does speak to the fear of some that volunteers threaten paid staff positions—in this instance, volunteers are almost literally looking to take the jobs of those they are working with.⁴⁷ Kevin B. Leonard writes that many archives volunteers are interns, recent graduates, or “otherwise

⁴⁵ “Volunteer Programs,” Maryland State Archives, January 27, 2016, accessed March 5, 2016, <http://msa.maryland.gov/msa/refserv/html/volunprogram.html>.

⁴⁶ “Behind-the-Scenes Volunteer Program,” Smithsonian, accessed March 5, 2016, <http://www.si.edu/Volunteer/Behind-the-Scenes-Volunteer>.

⁴⁷ Kevin B. Leonard, “Volunteers in Archives: Free Labor, But Not Without Cost,” *Journal of Library Administration* 52, no. 3/4 (April 2012): 314, accessed April 7, 2014, doi:10.1080/01930826.2012.684529; Frevert, “Archives Volunteers,” 152-153.

employable people, willing to fill time in an era of diminished employment prospects.”⁴⁸

There is a chance, however, that Leonard may be misjudging the motivations of those volunteers he sees as “filling time” before finding another position—they may also be volunteering to avoid a large resume gap, which could lessen their future employability. Uma Doraiswamy, a library and information science graduate, wrote an article with tips for other recent library science graduates looking for work. Her first tip was to volunteer in a library while searching for employment in an attempt to add skills to one’s resume.⁴⁹

CONCLUSION

Altruism and self-interest are extremely difficult to separate as motivations for volunteering, because even if individuals volunteer because they find the activity enjoyable, the enjoyment they feel is inherently a benefit. Therefore, it is perhaps best to define volunteering as a leisurely and pleasurable activity that is motivated by a combination of self-interest and altruism. Volunteering out of self-interest does not have to be a problem; it is in fact a natural motivator for most volunteers. Most volunteers working for the National Archives are motivated either by the social aspect of meeting new people and seeing friends regularly or by their interest in the historic record. Thanks to this, NARA has amassed a large group of dedicated volunteers whom are entrusted with a wide variety of tasks, some of which are complex and difficult. This thesis will

⁴⁸ Leonard, “Volunteers in Archives,” 314.

⁴⁹ Uma Doraiswamy, “Tips for Library and Information Science Students Seeking Employment and Entering the Workforce,” *Collaborative Librarianship* 3, no. 3 (July 2011): 176, accessed February 19, 2016, <https://ezproxy.mtsu.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=ofm&AN=67046968&site=eds-live&scope=site>.

primarily examine volunteers who find their work leisurely and are motivated by the pleasure they receive from the work. Defining volunteering and motivation is an important step in understanding crowdsourcing, the history of which is discussed in the next chapter.

CHAPTER III:

HISTORY OF CROWDSOURCING IN ACADEMIC INSTITUTIONS

In recent years, crowdsourcing, open source software, and websites filled completely by user-generated content seem to have all but taken over the Internet. But what exactly is crowdsourcing, and where did it come from? What potential consequences does Web 2.0 have for society? These questions are surprisingly difficult to answer, and those who try come up with differing answers. Every article seems to contain a different definition for crowdsourcing, and scholars cannot agree whether the interconnectedness of the Internet is a boon or a hindrance to modern society and future generations. All of this confusion, however, is simply because crowdsourcing is such a new phenomenon and its uses are constantly changing. It is possible to determine a usable definition of crowdsourcing and to trace a basic history of the idea both generally and specifically as it is used in archives by examining some popular projects and publications.¹

¹ Because crowdsourcing is a new phenomenon, historians must be creative in their use of sources when studying it. There are few historical primary sources examining its development. This paper examines a large number of unorthodox sources such as blog posts and tech journals, because these were the only truly primary sources discussing many crowdsourced projects. Most scholarly articles that have been written on crowdsourcing are case studies, and those are used here as primary sources as well. There is somewhat more historiography on open source, which will be addressed later in this paper. I chose to write the section on “Internet History” in reverse chronological order, because my research started at the present day and moved backwards, and because authors seemed to clearly be looking back at what had come before to explain what was currently happening. My list of crowdsourced projects, however, is organized chronologically for two reasons: first, to bring the reader back into the present, and second, in an attempt to get a true sense of “what happened when” and to illustrate that crowdsourcing has a strong history in academia.

DEFINING CROWDSOURCING

Defining an idea that is still being formed is not an easy task. For a preliminary definition, as well as to attempt to begin a search of the history of the phenomenon, it was helpful to consult the Oxford English Dictionary. Oxford defined crowdsourcing as “the practice of obtaining information or services by soliciting input from a large number of people, typically via the Internet and often without offering compensation.”² Oxford noted that the word was first used by journalist Jeff Howe in an article in *Wired* magazine, which was published in June of 2006 but was available to read in late May. (Because of this, other sources responding to his piece were published with earlier publication dates.)³

Even Howe provided two different definitions for crowdsourcing. On his website discussing the topic, he listed one definition that he considered more appropriate for written use, and a second, short definition that he called his “soundbyte version.” The first defined crowdsourcing as taking a task which would normally be performed by a specific person or group and instead “outsourcing it to an undefined, generally large group of people.”⁴ Under this definition, then, crowdsourcing was literally outsourcing a task to a crowd. Howe’s second definition was similar, but required some knowledge of

² “Crowdsourcing, N.,” OED Online (Oxford University Press), accessed February 19, 2016, <http://www.oed.com.ezproxy.mtsu.edu/view/Entry/376403#eid288590739>.

³ Ibid.

⁴ Jeff Howe, “Crowdsourcing,” accessed October 20, 2013, <http://www.crowdsourcing.com/cs/>.

the Internet and computing: he wrote that crowdsourcing was “the application of Open Source principles to fields outside of software.”⁵ At its most basic level, open source software is software in which the code is shared openly, with any interested individual allowed to improve upon and modify the code. The resulting programs are free for public use.⁶

A more recent article, “Towards an Integrated Crowdsourcing Definition,” by Enrique Estellés-Arolas and Fernando González-Ladrón-de-Guevara, noted the wide variance in crowdsourcing definitions and attempted to create one definition of crowdsourcing that could be applied to all endeavors. The authors noted that the likely reason for so much variance among articles and authors was because crowdsourcing was such a new idea that organizations were constantly coming up with new projects and new ways to implement crowdsourcing. The idea was therefore constantly evolving and so far this evolution has made it difficult to decide upon a single definition.⁷ In an attempt to reach a conclusive definition of crowdsourcing, the authors surveyed an impressive breadth of literature on the subject to determine the working definition in each article and book. They integrated these interpretations together into a single definition that they

⁵ Jeff Howe, “Crowdsourcing,” accessed October 20, 2013, <http://www.crowdsourcing.com/cs/>.

⁶ “The Open Source Definition (Annotated),” Open Source Initiative, accessed November 4, 2013, <http://opensource.org/osd-annotated>.

⁷ Enrique Estellés-Arolas and Fernando González-Ladrón-de-Guevara, “Towards an Integrated Crowdsourcing Definition,” *Journal of Information Science* 38, no. 2 (April 2012): 198, accessed November 3, 2013, doi: 10.1177/0165551512437638.

argued effectively represented the gamut of crowdsourcing initiatives currently being used:

Crowdsourcing is a type of participative online activity in which an individual, an institution, a non-profit organization, or company proposes to a group of individuals of varying knowledge, heterogeneity, and number, via a flexible open call, the voluntary undertaking of a task. The under-taking of the task, of variable complexity and modularity and in which the crowd should participate bringing their work, money, knowledge, and/or experience, always entails mutual benefit. The user will receive the satisfaction of a given type of need, be it economic, social recognition, self-esteem, or the development of individual skills, while the crowdsourcer will obtain and utilize to their advantage what the user had brought to the venture, whose form will depend on the type of activity undertaken.⁸

As part of their definition, Estellés-Arolas and González-Ladrón-de-Guevara noted that all crowdsourcing initiatives had eight aspects in common, and therefore in order for a project to be considered crowdsourcing, it must have met all eight standards. Their standards are:

- a) there is a clearly defined crowd;
- b) there exists a task with a clear goal;
- c) the recompense received by the crowd is clear;
- d) the crowdsourcer is clearly identified;
- e) the compensation to be received by the crowdsourcer is clearly defined;
- f) it is an online assigned process of a participative type;
- g) it uses an open call of variable extent;
- h) it uses the internet.⁹

According to Estellés-Arolas and González-Ladrón-de-Guevara, deciding whether a project is crowdsourcing is simple: apply the eight standards listed above to a project. If it

⁸ Estellés-Arolas and González-Ladrón-de-Guevara, "Towards an Integrated Crowdsourcing Definition," 197.

⁹ Ibid., 198.

meets all requirements, it is crowdsourcing. If any of the requirements is not met, the project does not constitute crowdsourcing.

Estellés-Arolas and González-Ladrón-de-Guevara further clarified several aspects of their requirements to make it easier to evaluate crowdsourced projects. They defined the crowd as “a large group of individuals,” but the crowd’s numbers and skills could vary depending on the task.¹⁰ The problem that the crowd solved could be simple or complex, but had to be clearly defined. This is why most do not consider YouTube to be a crowdsourcing initiative—contributors are not solving a problem with a clear end result. The authors also departed from other definitions of crowdsourcing to specify that the crowd must receive some sort of payment. This could be monetary, however it was often simply the feeling of accomplishment that resulted from having contributed to public knowledge.¹¹ The task that the public solved could be given by almost anyone, be it a company, institution, or an individual, so long as they had the ability to monitor the project through to its end. In return for their efforts they received the answer to whatever problem they posed to the crowd. The crowdsourcer must have solicited help via some form of an open call, whether it was truly open or whether participants need some sort of basic skill. For example, participants in a crowdsourced movie would need to have basic film skills. Finally, also departing from other definitions, Estellés-Arolas and González-

¹⁰ Estellés-Arolas and González-Ladrón-de-Guevara, “Towards an integrated crowdsourcing definition,” 194.

¹¹ Ibid., 194-195.

Ladrón-de-Guevara argued that the crowdsourcing initiative had to take place over the Internet.¹²

This thesis will use the definition set by Estellés-Arolas and González-Ladrón-de-Guevara as a definition of crowdsourcing. By examining a variety of sources, the researchers have developed a comprehensive definition that can be applied easily to any venture to determine whether or not it qualifies as crowdsourcing. It takes into account the evolution of crowdsourcing over its history, and although crowdsourcing will likely continue to evolve and may outdate this definition, it is currently useful to identify projects.

THE HISTORY OF INTERNET CROWDSOURCING

On May 25, 2006, Nick Douglas wrote a post on his blog called “Wagged, sagged, body-bagged: Things we’ve decided are dead.” It featured a table of technology trends that were “wagged” (currently being discussed), sagged (old news), and body-bagged (really old news or dead). Douglas wrote that outsourcing was “sagged” and open source was “body-bagged,” but under the “wagged” column was the term “crowdsourcing,” linked back to Jeff Howe’s article in the June 2006 *Wired* magazine.¹³ Howe’s article was the first to put a name to crowdsourcing, although it was a phenomenon that had been taking place for quite some time.

¹² Ibid., 194-196.

¹³ Nick Douglas, “Wagged, sagged, and body-bagged: Things we’ve decided are dead,” *Valleywag [Gawker] Blog*, May 25, 2006, accessed October 19, 2013, <http://gawker.com/176180/wagged-sagged-body+bagged-things-weve-decided-are-dead>.

Howe's article "The Rise of Crowdsourcing" was an innovative piece describing the new practice in the business world. On his blog, Howe noted that when the article first came out, there were three Google results for the word "crowdsourcing;" about a week later there were 182,000.¹⁴ Howe argued that crowdsourcing was a result of the move towards open source software development, and refuted accusations of amateurism, saying "the open source software movement proved that a network of passionate, geeky volunteers could write code just as well as the highly paid developers at Microsoft."¹⁵ Thanks to the crowdsourcing initiative, individuals now had a more productive outlet for their hobbies, could be helpful to the greater public, and could sometimes make money from their hobbies.¹⁶

Crowdsourcing may be a new idea, but the concept that Howe said it is derived from, open source, had been evolving for years. In 2002, David Bretthaur wrote that the idea of open source began in the 1970s, when groups of computer programmers from a few labs, led by MIT, shared the code software from their lab programs to help others fix bugs in their own programs. The scientists considered themselves a part of a community and it made sense to them to help each other with the problems that they were having. Many open source proponents say that the system is similar to the scientific method in its constant sharing of ideas, problems, and results. In 1989, programmers at Berkeley

¹⁴ Jeff Howe, "Birth of a Meme," *Crowdsourcing.com*, May 27, 2006, accessed October 20, 2013, http://www.crowdsourcing.com/cs/2006/05/birth_of_a_meme.html.

¹⁵ Jeff Howe, "The Rise of Crowdsourcing," *Wired* 14 no. 6, (June 2006), accessed October 19, 2013, <http://www.wired.com.ezproxy.mtsu.edu/wired/archive/14.06/crowds.html>.

¹⁶ Ibid.

developed their own operating system and began releasing it to the public. Open source became the most popular, however, in 1991 when Linus Torvalds developed Linux and posted a message on a forum requesting help with the coding. Since then it has spread outside of operating systems. Netscape Navigator, the Internet browser, made their code public in 1998 and became known as Mozilla, the company that now runs Firefox and various other open source projects.¹⁷ Although many archival institutions use commercial software to host exhibits and collections, there are many examples of free open source software for the archival profession as well. One example is ArchivesSpace, an information management system born of a merger between prior open source software programs Archon and Archivist's Toolkit. Omeka is a program designed by the Roy Rosenzweig Center for History and New Media intended for designing and hosting exhibits and digital collections, and Scripto, a plug-in compatible with Omeka and also

¹⁷ This is a short overview of a few major milestones in the history of open source. For a more in-depth discussion, read David Bretthaur's "Open Source Software: A History," *Information Technology and Libraries* 21, no. 1 (March 2002): 4, accessed December 9, 2013,

<http://ezproxy.mtsu.edu:3443/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=aph&AN=6607909&site=eds-live&scope=site>, Chris DiBona, Sam Ockerman, and Mark Stone, editors, *Open Sources: Voices from the Open Source Revolution* (Sebastopol, CA: O'Reilly Media, Inc., 1999), and Chris Dibona, Mark Stone, and Daneese Cooper, *Open Sources 2.0: The Continuing Evolution* (Sebastopol, CA: O'Reilly Media, Inc., 2005). Bretthaur argues that open source differs from shareware and freeware because of differences in access to the code and copyright restrictions, therefore this paper will only discuss open source. Bretthaur, "Open Source Software: A History," 4; DiBona, Ockerman, and Stone, introduction to *Open Sources: Voices from the Open Source Revolution*, 6-7; Bretthaur, "Open Source Software," 5-7; John C. Dvorak, "Upstarts Attack Microsoft Slackers," *PCMagazine* 21, no. 22 (December 24, 2002): 63, accessed December 9, 2013, <http://ehis.ebscohost.com.ezproxy.mtsu.edu/eds/pdfviewer/pdfviewer?sid=a1a05f93-a5f1-4e3c-8659-61a1c509dd02%40sessionmgr115&vid=2&hid=101>.

designed by the Roy Rosenzweig Center, provides a platform for crowdsourcing.¹⁸ Open source software has provided an opportunity for individuals to utilize their skills to develop software that is not only low cost but is also customizable to the needs of the individual or institution using it.

As part of his article, Howe listed a series of crowdsourcing initiatives that were already underway. His most beneficial example for those volunteering was InnoCentive, a scientific and industrial crowdsourcing initiative where “seekers” (usually companies or organizations) paid “solvers” to find answers to problems that companies had failed to solve. Solvers received payments between \$10,000 and \$100,000 per problem solved. Most solvers were either undergraduates in science programs, or people working out of their garage. Scholars who have studied the site noted that those who found workable solutions to problems tended not to have any professional training in the science of the problem.¹⁹

Crowdsourcing does not benefit everyone. One professional photographer who supplied stock photography was losing clients, only to discover that former customers were buying images from sites like iStockphoto, a crowdsourced stock photography site where photos cost between \$1-\$5. Even large companies bought images from the site,

¹⁸ “ArchivesSpace Home,” ArchivesSpace, accessed February 20, 2016, <http://www.archivesspace.org/>; “Archon: The Simple Archival Information System,” Archon, accessed February 20, 2016, <http://www.archon.org/>; “Archivists’ Toolkit,” Archivist’s Toolkit, accessed February 20, 2016, <http://www.archiviststoolkit.org/>; “Omeka,” Omeka, accessed February 20, 2016, <http://omeka.org/>; “Scripto,” Scripto, accessed February 20, 2016, <http://scripto.org/>.

¹⁹ Howe, “The Rise of Crowdsourcing,” also see InnoCentive’s website, <http://www.innocentive.com/>.

and it was so successful that it was bought by Getty Images. Another crowdsourcing site known as Amazon Mechanical Turk allowed companies to post simple tasks online for people to complete, and usually paid a few cents or a few dollars for each finished task. Howe interviewed one company who posted the task of writing flowcharts. The company needed workers who knew Java and Microsoft, and found workers on Mechanical Turk, some of whom had quit jobs in software development to raise families. Because coding ability involved a high level of skill, the company paid workers \$5 per task, a much higher price than most other tasks on Mechanical Turk. As a comparison, however, the company would have paid \$2,000 to outsource the tasks to professionals.²⁰ As those scholars and authors who oppose crowdsourcing argue, (these concerns will be discussed in greater detail later in this chapter), this is one of the downsides of crowdsourcing: in an effort to reduce their bottom line, companies risk turning websites into Internet sweatshops. When used effectively, however, crowdsourcing benefits the company and the individual.

Howe also provided “5 Rules of the New Labor Pool,” which were his tips for utilizing crowdsourcing to its fullest. The first rule was that the crowd is all around the world, so the job must be available wherever they are. This is why the Internet is an important aspect of the definition of crowdsourcing. In the twenty-first century, it is the easiest way to work from home. The second rule was to recognize that the people in the crowd were doing the work in their spare time, which they did not have much of, so tasks should not take too long. His third rule refuted the arguments of many of those who

²⁰ Howe, “The Rise of Crowdsourcing.”

opposed crowdsourcing: “The crowd is full of specialists.”²¹ As the company looking for flowcharts discovered, people may leave jobs for a variety of reasons, and in the current economy there are many people who are under or unemployed and, hence, willing to apply their specialist skills to a crowdsourcing project. Howe’s fourth rule was more in keeping with the fears of dissenters, stating that most of the results of crowdsourcing will be awful and there will need to be lots of sorting to find the proverbial needle in the haystack. Finally, Howe’s last rule was that “the crowd finds the best stuff.” Allowing the crowd to sort through data will result in them fixing errors and reporting whatever they find interesting or funny.²² A good example is “likes” on YouTube videos. There are too many videos on YouTube to ever possibly view, but the crowd watches them and recommends videos by commenting on and liking them. Other viewers can then learn which videos they might find interesting or worth viewing.

Crowdsourcing had actually been happening for quite some time before Howe wrote his now-famous article. In January of 2005, before Howe wrote his article, Richard Watson wrote a blog in which he also discussed the idea of “open source” and how it had spread beyond the scope of software development. Watson did not use the term *crowdsource* (Howe was the first), but it is clear that they were discussing the same idea. Watson argued that open source principles worked well and worked quickly, and noted that “the idea has been transferred” to Wikipedia, “aeroplane design, cola recipes, film

²¹ Jeff Howe, “5 Rules of the New Labor Pool,” *Wired* 14 no. 6 (June 2006), accessed October 20, 2013, <http://www.wired.com.ezproxy.mtsu.edu/wired/archive/14.06/labor.html>.

²² Howe, “5 Rules of the New Labor Pool.”

scripts, and beer.”²³ Critics had claimed that open source was really just a giant test group, but Watson argued that this was not the case for several reasons: first, because of the sheer size involved, second, because people had a vested interest in the project, and finally, because the nature of crowdsourcing groups and test groups is fundamentally different. In crowdsourcing and open source initiatives, people are not testing products, they are suggesting products: “focus groups usually ask people to react to ideas. Open source asks people for solutions and allows ideas to build cumulatively.”²⁴ This difference is key, and crowdsourcing initiatives not only allow ideas to grow for the company, but also for the individuals assisting with the problems. For example, Proctor and Gamble had a goal of fifty percent of their new ideas coming from open source.²⁵ Watson’s article was important because it truly stressed and inspired collaboration between businesses and the public. He also succinctly phrased the difference between crowdsourcing and open source concepts and test groups.

In 2003, Thomas Goetz also documented the history of the open source movement. Open source as we view it started in 1991 with the Linux operating system. Goetz argued that coders were the first to transition open source to computers simply because they had the knowledge to do so. In 1991, Internet usage was not yet widespread

²³ Richard Watson, “Tech@Work: A problem shared is a problem solved,” *The Edge Singapore*, January 24, 2005, http://www.lexisnexis.com.ezproxy.mtsu.edu/lnacui2api/results/docview/docview.do?start=9&sort=BOOLEAN&format=GNBFI&risb=21_T18420586719.

²⁴ Watson, “Tech@Work.”

²⁵ Ibid.

enough for the general public to know how to utilize it to its full advantage, but coders were better aware of how to use it. In recent years, open source has become more popular as Internet literacy has spread, and its collaboration power helped people who were struggling to work under intellectual property laws.²⁶

Goetz pointed to Wikipedia as an example of open source. In 1999, founder Jimmy Wales first attempted a similar type of encyclopedia called “Nupedia,” where users had to apply to contribute articles. The articles were then peer-reviewed, similar to an academic journal. The encyclopedia was not especially popular. He tried again in 2001 with Wikipedia, allowing anyone to add information and make edits, and the site was much more successful. Some people caused problems, and Wikipedia responded by locking some pages and banning certain contributors. There were many reasons that contributors chose to participate – they may have been obsessive about fixing mistakes they saw; they may have wanted to show off how smart they were; they may have felt they were helping the world; and they may simply have enjoyed the chance to utilize their knowledge.²⁷ As per the definition, the crowd was being rewarded for their efforts, not monetarily, but in whatever way they personally felt they needed to be rewarded.²⁸

²⁶ Thomas Goetz, “Open Source Everywhere,” *Wired* 11 no. 11 (November 2003), accessed October 20, 2013, http://www.wired.com.ezproxy.mtsu.edu/wired/archive/11.11/opensource_pr.html.

²⁷ Ibid.

²⁸ For more information on the history of Wikipedia, the motivations of its volunteers, and its successes, see Hoda Batiyeh and Jay Pfaffman, “Volunteers in Wikipedia: Why the Community Matters,” *Journal of Educational Technology and Society* 13 no. 2 (July 2010), accessed February 2, 2016, <https://ezproxy.mtsu.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&d>

In his blog post “Crowdsourcing: A Definition,” Howe wrote that the definition of crowdsourcing was being changed by the crowd itself to match Yochai Benkler’s theory of commons-based peer production.²⁹ Benkler defined commons-based peer production as “large aggregations of individuals independently scouring their information environment in search of opportunities to be creative in small or large increments. These individuals then self-identify for tasks and perform them for a variety of motivational reasons.”³⁰ He introduced this theory in his 2002 article “Coase’s Penguin, or, Linux and the Nature of the Firm.” Similar to other writers, Benkler described the rise of production via large groups, rather than single companies. He noted that scientific research, and academic work in general, was the primary example of commons-based peer production because everyone contributed what they knew, peer reviewed each other’s work, and even had informal discussions concerning research.³¹

[baph&AN=52045409&site=eds-live&scope=site](https://ezproxy.mtsu.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=baph&AN=52045409&site=eds-live&scope=site); Sam Ransbotham and Gerald C. Kane, “Membership Turnover and Collaboration Success in Online Communities: Explaining Rises and Falls from Grace in Wikipedia,” *MIS Quarterly* 35 no. 3 (September 2011), accessed February 2, 2016, <https://ezproxy.mtsu.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=bth&AN=63604897&site=eds-live&scope=site>.

²⁹ Jeff Howe, “Crowdsourcing: A Definition,” *Crowdsourcing.com Blog*, June 2, 2006, accessed October 20, 2013, http://www.crowdsourcing.com/cs/2006/06/crowdsourcing_a.html.

³⁰ Yochai Benkler, “Coase’s Penguin, or, Linux and the Nature of the Firm,” *The Yale Law Journal* 112 no. 3 (December 1, 2002): 376, accessed October 20, 2013, <http://www.jstor.org/stable/1562247>.

³¹ *Ibid.*, 375.

In his article Benkler explained his theory of commons-based peer production. In this theory, volunteers chose which tasks they wanted to perform, and volunteers had many different reasons for choosing to participate. Their motivation was often psychological, which Benkler argued was generally all the motivation necessary since tasks were small and did not take long to perform. There were three common aspects of successful projects. First, successful projects could be split into small projects that did not rely on each other to be understood, so that volunteers could complete as much or as little as they wanted. Second, successful projects needed to be small so that large numbers of people would be willing to participate. Finally, in successful projects the cost of quality control and putting the small projects back together into the finished project must be low enough to not sink the project entirely.³² The final point was important because it was difficult – the labor involved in quality control could be intensive. As such, most successful projects had some sort of peer review system. For example, the website Slashdot was a user-generated news source where users posted links to stories on other websites along with comments on the articles. The website did not check for the accuracy of the posts, but allowed users to post comments arguing the validity of information presented. Benkler compared this to a system of peer review.³³

This sounds like the worst nightmare of those arguing that the public does not have the credentials necessary to disseminate accurate information, but Benkler argued that commons-based peer production did not replace companies or traditional methods of

³² Benkler, “Coase’s Penguin,” 375-379.

³³ Ibid., 376, 393-394.

production. Nor did he believe that commons-based peer production was always the better option. Benkler's argument was simply that it was a different model and that it inherently "has certain systematic advantages...in identifying and allocating human capital/creativity."³⁴ These advantages existed for two reasons. First, people were allowed to choose for themselves which tasks they performed. As Benkler saw it, individuals are their own best judge of their talents, interests, and resources. Secondly, larger groups had more access to more resources and could better transmit these resources and ideas.³⁵ It worked as a six degrees of separation for information – everyone knows something, so widening the crowd widened the breadth of information available.

CROWDSOURCED PROJECTS

A brief history of the variety of crowdsourced projects shows that crowdsourcing has for the most part proved successful. A variety of academic fields have embraced it, and the results they have achieved have been impressive. While crowdsourcing may sometimes have its drawbacks, the results of the projects at hand show that it is worth the time and effort it entails.

Although according to the scale devised by Estellés-Arolas and González-Ladrón-de-Guevara Wikipedia does not meet all of the criteria of crowdsourcing,³⁶ it is still relevant to the history of crowdsourcing because it was an early project that has gained

³⁴ Benkler, "Coase's Penguin," 381.

³⁵ Ibid., 376-377.

³⁶ Estellés-Arolas and González-Ladrón-de-Guevara, "Towards an integrated crowdsourcing definition," 197.

much success. “Wikis” were first invented by Ward Cunningham for use on his website WikiWikiWeb and named after the Hawaiian word wikiwiki, which means quickly. While the computer definition of a wiki is likely fairly common knowledge today, an article in 2003 had to define it, noting that it was a “hypertext document that allows its pages to be quickly and easily edited or deleted by any visitor.”³⁷

When one author reviewed Wikipedia for his article “Site of the Week” in 2003, he noted that it had over 130,000 articles. And yet, none of his searches returned any terrible articles or evidence of vandalism. The author attributed this to the idea that many contributors on Wikipedia were “people writing on subjects near to their hearts,” and would therefore not allow vandalism or poor writing to stay online.³⁸ Wikipedia was still new enough that it had just over 2,000 websites linking back to it, and as of 2003 Cunningham’s site WikiWikiWeb was still more popular.³⁹ As of this writing, however, Wikipedia has now grown to contain 5,070,861 articles written in English.⁴⁰

³⁷ Kinley Levack, “If Two Heads Are Better than One, Try 7,000 With Wikipedia,” *EContent* 26 no. 4 (April 2003): 12, accessed November 1, 2013, <https://ezproxy.mtsu.edu:3443/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=llf&AN=501020266&site=eds-live&scope=site>; Sean Carroll, “Site of the Week: Wikipedia,” *PCMagazine.com* (June 6, 2003) accessed November 1, 2013, <http://www.lexisnexis.com.ezproxy.mtsu.edu/hottopics/lnacademic/>.

³⁸ Carroll, “Site of the Week: Wikipedia.”

³⁹ Carroll, “Site of the Week: Wikipedia”; Levack, “If Two Heads are Better than One,” 12.

⁴⁰ “Main Page,” Wikipedia, the free encyclopedia, accessed February 2, 2016, http://en.wikipedia.org/wiki/Main_Page.

Probably the oldest example of crowdsourcing in a form that relates to archives is the FamilySearch indexing project, which began in August of 2005. FamilySearch is a free, Web-based genealogy service provided by the Church of Jesus Christ of Latter-day Saints as part of their belief system that “families are forever,” and that researching genealogy is a way of connecting with ancestors.⁴¹ The service is free anyone who wishes to use it, regardless of religious belief. Records and accounts are free to access. There are 4,500 physical research centers located around the world, and the website offers twenty-four hour support via telephone and chat. The indexing project enlists volunteers to transcribe names, dates, and other vital information from documents in order to make them searchable. The records for the project come from all around the world, and as of this writing, volunteers had indexed 1,281,531,734 records.⁴²

Other academic crowdsourcing groups took a few years to catch up to the FamilySearch’s project. GalaxyZoo, a crowdsourcing for astronomy hosted on the crowdsourcing platform Zooniverse, was founded in July of 2007 with the goal of categorizing different types of galaxies. Volunteers classify galaxies by looking at an image of a galaxy, then answering a series of questions about the picture with drawings to help (see figure 1). For example, one question might ask, “Is the galaxy simply smooth and rounded, with no signs of a disk?” Below the questions are drawings of a round

⁴¹ Rose Holley, “Crowdsourcing: How and Why Should Libraries Do It?” *D-Lib Magazine* 16 no. 3/4 (March/April 2010), accessed November 3, 2013, doi: 10.1045/march2010-holley; “About FamilySearch,” FamilySearch.org, accessed November 3, 2013, <https://familysearch.org/about>.

⁴² “About FamilySearch”; Familysearch.org, *FamilySearch Indexing: How It Works*, <https://familysearch.org/indexing/>; “Indexing Overview,” FamilySearch.org, accessed February 2, 2016, <https://familysearch.org/indexing/>.

galaxy and galaxies with disks to provide an example for volunteers and help them choose which type of galaxy they are viewing.⁴³

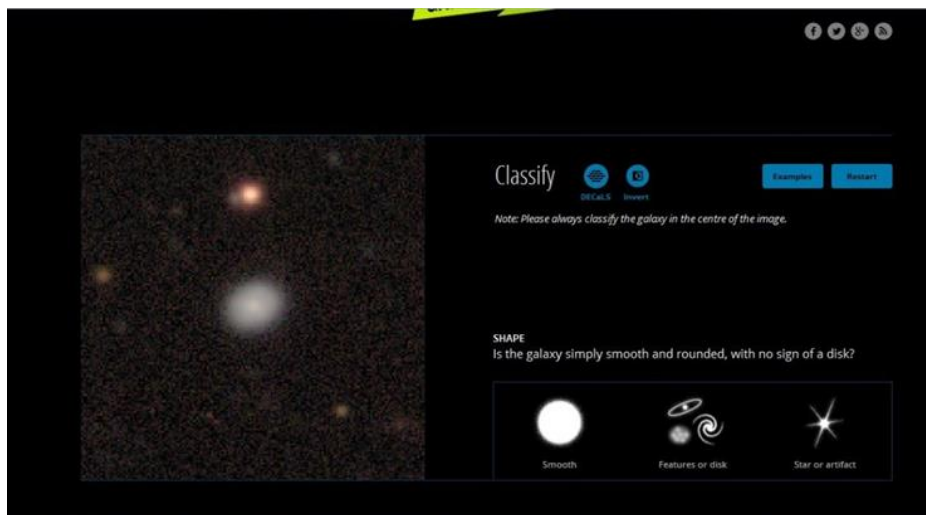


Figure 1. GalaxyZoo identification platform. Image courtesy GalaxyZoo/Zooniverse. GalaxyZoo is a Zooniverse.org project.

The following summer Australia's archives joined the Australian Newspapers Digitisation Program. This program is different from other archival ventures because rather than asking volunteers to transcribe records, the archivists asked them to correct mistakes in records that had already been transcribed via Optical Character Recognition.

⁴³ Holley, "Crowdsourcing: How and Why Should Libraries Do It?"; "Classify," GalaxyZoo, accessed December 9, 2013, <http://www.galaxyzoo.org/#/classify>.

Users can also comment on newspapers and add tags. The project is still in existence and now operates under the name “Trove.”⁴⁴

Less than a month later, on September 2, 2008, the genealogy website Ancestry.com launched its own version of a records indexing project known as the World Archives Project. They encouraged archives to donate digital scans of their materials to Ancestry, and in return, the archives would receive a digital copy of the index when it was completed. Similar to the FamilySearch project, Ancestry is not looking to transcribe entire documents, but to merely provide indexing to make the documents searchable. Ancestry posts the digital images to their website and volunteers transcribe information such as names, dates, and locations. This way, researchers on Ancestry.com can type in their relative’s name and be able to find the record quickly, rather than sifting through piles of information. In keeping with the definition of crowdsourcing, Ancestry rewards volunteers with a discount on renewals of their membership to the website.⁴⁵

Many crowdsourcing initiatives point to Transcribe Bentham as a top example of crowdsourced archival transcription, because it is a large and well-organized project. The Bentham Project at University College London (UCL) originally formed in 1958 to transcribe The Collected Works of Jeremy Bentham and publish them in a series of books. Prior to beginning their crowdsourcing initiative, they had only produced twenty-

⁴⁴ Holley, “Crowdsourcing: How and Why Should Libraries Do It?”; “Trove,” National Library of Australia, accessed December 9, 2013, <http://trove.nla.gov.au/>.

⁴⁵ “Ancestry.com Announces the World Archives Project,” *Ancestry.com Blog*, September 4, 2008, accessed November 3, 2013, <http://blogs.ancestry.com/ancestry/2008/09/04/ancestrycom-announces-the-world-archives-project/>; “About the Ancestry.com World Archives Project,” Ancestry.com, accessed November 3, 2013, <http://landing.ancestry.com/wap/learnmore.aspx>.

nine of what they expected to be seventy books. UCL began crowdsourcing in September of 2010. The project is considered to be “one of the first to try crowd-sourced [sic] transcription and to open up a traditionally rarified scholarly endeavor to the public.”⁴⁶ Volunteers work from the Transcribe Bentham “Transcription Desk” to first enter transcriptions of a manuscript into a textbox and then format the transcription with line breaks, page breaks, and other basic formatting, including marking text that is unreadable (see figure 2). University College London uses the transcriptions both for the eventual publishing of Jeremy Bentham’s works, but also to post them online as digital archives.⁴⁷ The Bentham Project notes that they looked to GalaxyZoo and the Australian Newspapers Digitisation Program for inspiration. During the six-month pilot phase, volunteers completed 569 pages of transcriptions and started on an additional 440 pages.⁴⁸

⁴⁶ Tim Causer, Justin Tonra, and Valerie Wallace, “Transcription maximized, expense minimized? Crowdsourcing and editing The Collected Works of Jeremy Bentham,” *Literary and Linguistic Computing* 27 no. 2, (June 2012): 120, accessed November 3, 2013, <https://ezproxy.mtsu.edu:3443/login?url=http://search.ebscohost.com.ezproxy.mtsu.edu/login.aspx?direct=true&db=edswah&AN=0003041999000001&site=eds-live&scope=site>; Patricia Cohen, “For Bentham and Others, Scholars Enlist Public to Transcribe Papers,” *New York Times*, December 27, 2010, accessed October 20, 2013, <http://www.nytimes.com/2010/12/28/books/28transcribe.html>.

⁴⁷ “Transcribe Bentham: Transcription Desk,” University College London, accessed December 8, 2013, http://www.transcribe-bentham.da.ulcc.ac.uk/td/Transcribe_Bentham.

⁴⁸ Causer, Tonra, and Wallace, “Transcription maximized, expense minimized?” 120, 125.

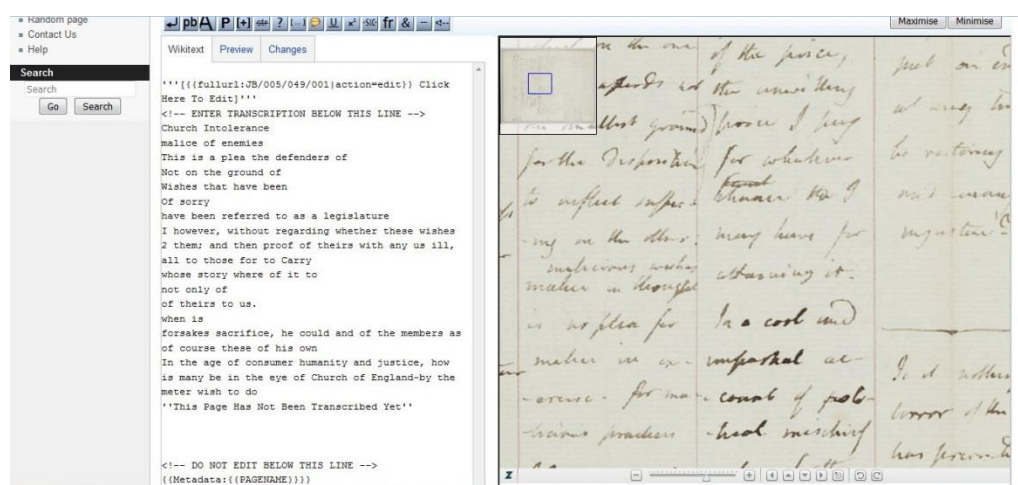


Figure 2. Transcribe Bentham Transcription Desk. Image courtesy of UCL Special Collections.

Two other archives began crowdsourcing projects in the spring of 2011. The New York Public Libraries began a project known as “What’s on the Menu?” to transcribe historic menus from New York restaurants. They have menus dating to the 1840s, and are hoping to make the text searchable. As of this writing, the library has transcribed 17,545 menus. They also list another bonus that archives receive from crowdsourcing projects: they hope the publicity surrounding the project will help researchers see the significance of the collection and inspire them to use it.⁴⁹ About the same time as “What’s on the Menu?” was getting off the ground, the University of Iowa began its Civil War Diaries Transcription Project. The project began with Civil War-era diaries hosted on a digital platform for volunteers to transcribe. Archivists involved in the project noted that interest

⁴⁹ “What’s on the Menu?” NYPL Labs, New York Public Library, accessed February 2, 2016, <http://menus.nypl.org/>. “What’s on the Menu? About,” NYPL Labs, New York Public Library, accessed October 21, 2013, <http://menus.nypl.org/about/>;

in the project was initially low for the first two months, until someone posted a link to their site on Reddit, at which point they received so many visitors that their server crashed. Iowa's project was so popular that they developed a larger project, called DIY History (figure 3,) that offers Civil War papers, women's diaries, correspondence relating to the Transcontinental Railroad, and historical cookbooks for public transcription.⁵⁰ Transcribe Bentham, What's on the Menu? and DIY History are all currently ongoing.

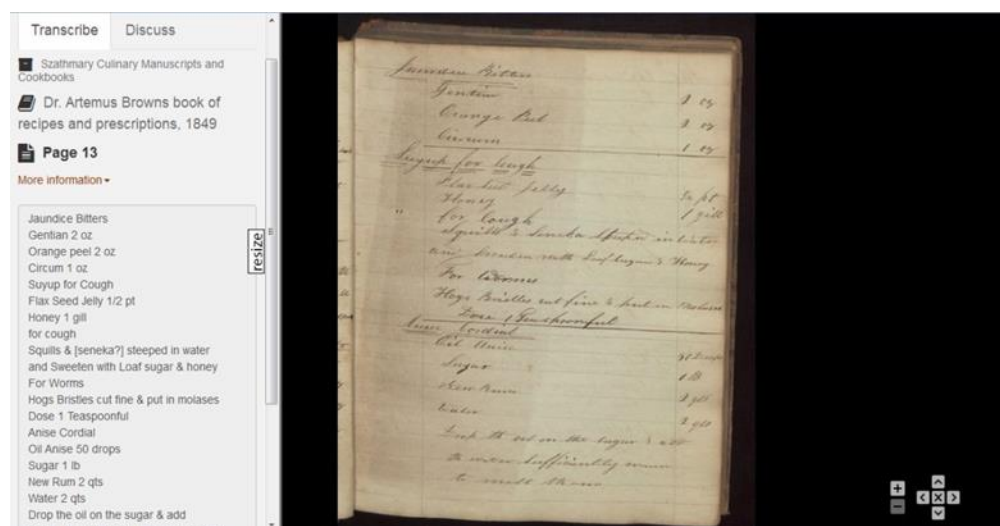


Figure 3. DIY History transcription platform. Image courtesy of University of Iowa Libraries.

⁵⁰ Nicole Saylor and Jen Wolfe, "Experimenting with Strategies for Crowdsourcing Manuscript Transcription," *Research Library Issues* no. 277 (December 2011): 13, accessed October 21, 2013, <http://publications.arl.org/rli277/10>, "DIY History," The University of Iowa Libraries, accessed December 9, 2013, <http://diyhistory.lib.uiowa.edu/>.

There are several examples of other crowdsourcing projects that are currently being planned or are in use by a variety of academic institutions. A few examples of this include New York Times Madison, the Indigenous Digital Archive, and the Crowd Consortium. New York Times Madison allows volunteers to find, tag, and transcribe advertisements from within digitized copies of historic New York Times newspapers. The project is meant to help researchers of economic and cultural history to easily search for relevant advertisements. It was started by the Research and Development Lab at the New York Times, who recognized the research potential of the advertisements and developed new open source crowdsourcing software they named Hive, which is available for free download.⁵¹ The Indigenous Digital Archive is a project being planned in conjunction with the Museum of Indian Arts and Cultural. According to the September 2014 introductory blog post, the project seeks to gather documents relating to Native American history from a variety of other archival sources and host them on one website to assist researchers and genealogists who previously had to search nationwide for records. Volunteers will then be able to tag the documents so they are further searchable by names or topics.⁵² In addition, dozens of libraries, archives, museums, and other academic groups have joined together to create the Crowd Consortium, a group dedicated to the study of crowdsourcing. The group's mission is to "explore the potential for crowdsourcing for enhancing research, collections, and other aspects of their

⁵¹ "About Madison," The New York Times Madison, accessed March 14, 2016, <http://madison.nytimes.com/contribute>.

⁵² Naruto-Moya, "Welcome! Help Build the Indigenous Digital Archive," Indigenous Digital Archive, September 17, 2014, accessed March 13, 2016, <http://blog.indigenousdigitalarchive.org/post/2014/09/17/first>.

institutions.” The website provides readings, research, case studies, webinars, and events to help institutions using crowdsourcing collaborate to make their platforms better. The site also provides a list of current and potential crowdsourcing initiatives.⁵³

BUT WHAT HAPPENED TO THE GATEKEEPER?

As mentioned earlier, not all scholars have been comfortable with the idea of recruiting the public to do the work of professionals. Scholars dislike crowdsourcing and the nature of user-generated websites for several reasons. Some are afraid that underqualified amateurs are unable to adequately perform the work that would normally be entrusted to a professional. Others cite ethical or economic concerns. For example, responding to Jeff Howe’s article introducing the idea of crowdsourcing, Nick Douglas, author of the media blog Valleywag wrote a post titled “That’s not slave labor, that’s crowdsourcing!” He called crowdsourcing “an idea as old as serfdom,” reporting “‘Unskilled labor’ gets a makeover!”⁵⁴ As hyperbolic as it may seem to relate a voluntary act to slavery and medieval serfdom, Douglas was not alone in his distrust of public generated media.

In his 2008 book *The Cult of the Amateur*, Andrew Keen related his story of how he came to mistrust Web 2.0 (interactive Internet websites, such as social media and

⁵³ Mary Flanagan, “Who We Are,” Crowd Consortium, accessed March 14, 2016, <http://www.crowdconsortium.org/who-we-are-2/>.

⁵⁴ Nick Douglas, “Job market news: That’s not slave labor, that’s crowdsourcing!” *Valleywag [Gawker] Blog*, May 25, 2006, accessed February 6, 2016, http://www.lexisnexis.com.ezproxy.mtsu.edu/lnacui2api/results/docview/docview.do?star t=2&sort=BOOLEAN&format=GNBFI&risb=21_T18420586719.

blogging, as well websites where individuals could comment on articles, rather than simply read information presented) and the disastrous effects that he believed it would have on our society.⁵⁵ Keen told the story of how he first became disillusioned with Web 2.0. He was attending a camp run by Internet innovator Tim O'Reilly; the topic of the event was Web 2.0. At the time, Keen was involved in Internet music distribution, and he was trying to bring Bach and Bob Dylan to the attention of more people. Keen soon discovered that Web 2.0 was all about user-generated content and therefore involved people uploading their own music to websites for the review of others. Keen took issue with this model, arguing that he has always been more trusting of the work of professionals over amateurs, whether it be the work of doctors, lawyers, journalists, or

⁵⁵ Reviewers of Keen's work are divided over whether or not he is correct. For reviewers that agree with Keen, see Carol Tenopir, "Web 2.0: Our Cultural Downfall?" *Library Journal* 132, no. 20 (December 15, 2007): 36, accessed December 6, 2013, <http://ehis.ebscohost.com.ezproxy.mtsu.edu/eds/detail?sid=6b0c119e-38d1-496b-a98e-30b84b0a0171%40sessionmgr112&vid=1&hid=115&bdata=JnNpdGU9ZWRzLWxpdmUmc2NvcGU9c2l0ZQ%3d%3d#db=nyh&AN=27924988>; R.J. Stove, "The Cult of the Amateur: How Today's Internet is Killing our Culture (Critical Essay)," *National Observer* 74 (Spring 2007): 64, accessed December 9, 2013, <http://ehis.ebscohost.com.ezproxy.mtsu.edu/eds/detail?sid=c82553ea-47cb-40bf-889f-acbae554c55a%40sessionmgr110&vid=1&hid=115&bdata=JnNpdGU9ZWRzLWxpdmUmc2NvcGU9c2l0ZQ%3d%3d#db=edsgao&AN=edsgcl.173717050>; Reviewers who disagree with Keen include Kevin Keohane, "Unpopular Opinion. Everyone's an Expert on the Internet. Is That Such a Bad Thing?" *Communication World* 25 no. 1 (2008): 12, accessed December 9, 2013, <http://ehis.ebscohost.com.ezproxy.mtsu.edu/eds/detail?sid=c94f5dc7-6643-4c3b-990d-e761681bd6e0%40sessionmgr113&vid=1&hid=115&bdata=JnNpdGU9ZWRzLWxpdmUmc2NvcGU9c2l0ZQ%3d%3d#db=edsgao&AN=edsgcl.173021681>; David Harsanyi, "The Amateur's Hour: Is the Internet Destroying our Culture, or is it Just Annoying our Snobs?" *Reason* 39 no. 8 (Jan 2008): 66-68, accessed December 9, 2013, <http://ehis.ebscohost.com.ezproxy.mtsu.edu/eds/detail?sid=0f7bcb18-8c47-4fba-8b87-5926568835a4%40sessionmgr114&vid=1&hid=115&bdata=JnNpdGU9ZWRzLWxpdmUmc2NvcGU9c2l0ZQ%3d%3d#db=edsgao&AN=edsgcl.172291492>.

musicians.⁵⁶ What Keen failed to realize, however, is that Web 2.0 is still living up to his original goal of delivering “more music to more orifices.”⁵⁷ It may not be the music he intended, but thanks to websites with user-generated content, music aficionados now have access to a wider range of music than ever before. Keen’s choice of trusting professional doctors over someone with no credentials is not likely to receive much debate. However, the idea of what constitutes a professional is harder to define for professions such as musicians.

Much of the distrust of user-generated content stems from the assumption that the world is divided into experts and amateurs with nothing in between. For example, Keen assumes that individuals do not possess multiple talents. He questions whether people can ever achieve skill in anything if they divide their time amongst multiple activities, saying “In a world in which we are all amateurs, there are no experts.”⁵⁸ Keen’s bias is further revealed in the way he defines an amateur: a “hobbyist, knowledgeable or otherwise, someone who does not make a living from his or her field of interest, a layperson, lacking credentials, a dabbler.”⁵⁹ This is one of the weaknesses of Keen’s book – his inability to recognize that amateurs might indeed have something of value to offer. His definition does acknowledge that an amateur may be knowledgeable. But the most important aspect

⁵⁶ Andrew Keen, *The Cult of the Amateur: How Blogs, MySpace, YouTube, and the Rest of Today's User-Generated Media are Destroying Our Economy, Our Culture, and Our Values*, (New York: Random House, 2008), 12-14, xiii, 2,9.

⁵⁷ Ibid., 12-14.

⁵⁸ Keen, *The Cult of the Amateur* 38-39.

⁵⁹ Ibid., 36.

of the definition is that amateurs have no credentials and do not make a living practicing their craft. The opposite of amateur—someone who makes a living practicing a specific craft—is a professional. Keen, however, uses the terms “expert” and “professional” interchangeably throughout the work without acknowledging that the words have slightly different meanings. In actuality, a professional may not truly be an expert, but those who distrust Web 2.0 do not acknowledge this. Put differently, Keen assumes that individuals who are paid to perform a task are automatically more knowledgeable about that task than an unpaid enthusiast, which may not always be true.

Keen also takes issue with crowdsourcing (although he does not name it as such), which he seemed to find both naïve and deceptive. He describes a Wal-Mart commercial called “School My Way” that was crowdsourced from student contributions. He notes that companies crowdsourced work because of the economic advantage but also because consumers believed crowdsourced advertising was more realistic and aimed towards their needs. He finds this naïve and worries about the economic effects of crowdsourcing.⁶⁰ Keen and Douglas do make a good point in their mistrust of private sector companies making a profit off of underpaid—or unpaid—crowdsourcing volunteers. As they are in the public sector and are working to provide greater access to public goods, libraries and archives are often excluded from concern over volunteer exploitation. Companies in the private sector who use crowdsourced volunteers, however, raise real questions about poorly- or unpaid workers being used to raise profits. Ancestry.com is particularly questionable, given that at first glance, it appears to be similar to any other archives or

⁶⁰ Keen, *The Cult of the Amateur*, 61-62.

library, but it is in fact a privately held company making a profit. Therefore, those volunteers who are indexing records cannot later have access to the same records unless they choose to pay for a membership with the website.

Perhaps most of Keen's fears can be summed up in a single sentence on the relationship between experts and an unmediated Internet. Keen fears that on the new Internet "there are no gatekeepers to filter truth from fiction, genuine content from advertising, [or] legitimate information from outright deceit."⁶¹ Websites like Wikipedia, therefore, are dangerous because they "undermine" what students are taught in school by their teachers, who are qualified disseminators and gatekeepers.⁶² There will always be some who mistake satire for real news and look to blogs and other social media for their information. Keen takes a rather extreme stance, however, on the larger question of whether the public should be allowed access to all information or whether they should only be allowed to view what professionals have decided is true and appropriate for public consumption.

The idea of a "gatekeeper" is also present in archival science, where it is used to describe one theory of the role of the archivist in relation to reference and records. In this model, archivists often view themselves as holding extremely specialized knowledge, and therefore as being responsible for the records first and the needs of the researcher

⁶¹ Ibid., 65.

⁶² Ibid., ix.

second.⁶³ Archivists have come to realize, however, that this model often provides poor services to researchers. According to Catherine A. Johnson and Wendy A. Duff, many researchers feel intimidated by archivists, or as if they needed to prove themselves or develop a friendly relationship with the archivist in order to view the records they want to see. Johnson and Duff even quote historians who felt that they were “at [the archivists’] mercy” or that archivists were a “guard dog, and they were no help whatsoever.”⁶⁴ As a result, many archivists are moving away from this “gatekeeper” mentality to think of the researcher’s needs first and the needs of the records second. Kate Theimer discusses this in her article “What is the Meaning of Archives 2.0?”, where she provides examples of how “Archives 2.0” (what she considers to be a change in how archivists are thinking and practicing) differs from “Archives 1.0” (how archivists were previously performing their work). She argues that the new shift in archival science includes “archivists see[ing] their primary role as facilitating rather than controlling access” and that “today’s archivists understand their mission to be serving researchers, not records.”⁶⁵ This ideological shift makes sense, because as any archivist is quick to acknowledge, archives exist for research as well as preservation—so without someone to view the records and use them for research, there is little purpose in keeping the records in the first place.

⁶³ Luke J. Gilliland-Swetland, “The Provenance of a Profession: The Permanence of the Public Archives and Historical Manuscripts Traditions in American Archival History,” *The American Archivist* 54, no. 2 (1991): 163, 173, accessed February 19, 2016, <http://www.jstor.org/stable/40293549>.

⁶⁴ Catherine A. Johnson and Wendy M. Duff, “Chatting up the Archivist: Social Capital and the Archival Researcher,” *The American Archivist* 68, no. 1 (2005): 121-122, accessed February 19, 2016, <http://www.jstor.org/stable/40294259>.

⁶⁵ Kate Theimer, “What Is the Meaning of Archives 2.0?,” *The American Archivist* 74, no. 1 (2011): 61-62, accessed February 19, 2016, <http://www.jstor.org/stable/23079001>.

Another article, written shortly after the website's founding, criticized Wikipedia. However, the author (a librarian) used less extreme arguments and therefore seems less paranoid than Keen. In his column "Péter's Picks & Pans," Péter Jascó wrote that he believed the goals of Wikipedia were overzealous and addressed his fears of articles being plagiarized from other sources. He noted that Wikipedia was founded in January of 2001, and at the time of his writing approximately a year later, it had 16,000 articles with hopes of one day reaching over 100,000. He observed that a comparable encyclopedia, the Columbia Encyclopedia, only contained 51,682 articles. Jascó saw the goal of 100,000 articles as overzealous, not just because he doubted Wikipedia could reach those numbers. He argued that even if they did achieve such a large number of entries that would be excessive. He also noticed that many articles on Wikipedia were full of spelling, grammatical, and even factual mistakes. He considered Wikipedia to be "a joke at best," because the visible mistakes while navigating the website made it "look like a prank."⁶⁶ Jascó remarked that many of the articles on specific countries were taken word-for-word from the *CIA World Factbook*. He noted that even though this source was in the public domain, it still needed to be cited. Jascó doubted how well Wikipedia could survive and meet its goals. It is worth noting again that as of this writing, Wikipedia has

⁶⁶ Péter Jascó, "Péter's Picks & Pans," *Online Magazine* 26 no. 2 (March/April 2002): 81, accessed November 1, 2013, <https://ezproxy.mtsu.edu:3443/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=llf&AN=502875163&site=eds-live&scope=site>.

far surpassed their goal of 100,000 articles and has reached a total of over five million articles in English.⁶⁷

The early skeptics were not completely mistaken—crowdsourcing has its downsides. Not all projects are successful. And, even successful projects require careful planning and hard work. Before beginning a project, staff must confront the fear of loss of power from professionals to amateurs. Volunteers are bound to make mistakes (or, worse yet, may intentionally spam the site), and so crowdsourcing is therefore not entirely the free labor many expect it to be. It costs time and money to oversee such large pools of volunteers. Metadata generated through public tagging, where volunteers are presented with an image and asked to add any words or phrases they feel would make for quality search terms, may also consist of current slang or regional terms that may not remain useful in the future or may even be unrecognizable to users outside of a specific geographic location. Companies who offer “micropayments” (often consisting of a few cents for every task performed) also face ethical problems relating to fair pay for labor. However, cultural institutions that house public heritage and build relationships with volunteers are often excluded from such accusations.⁶⁸

What are some ways in which crowdsourcing projects can go wrong? Daniel Stowell, director and editor for the Papers of Abraham Lincoln, a project which seeks out

⁶⁷ Ibid., 82; “Main Page,” Wikipedia, the free encyclopedia, accessed February 2, 2016, http://en.wikipedia.org/wiki/Main_Page.

⁶⁸ Sally Ellis, “A History of Collaboration, a Future of Crowdsourcing: Positive Impacts of Cooperation on British Librarianship,” *Libri: International Journal of Library & Information Services*, 64 no. 1 (March 2014): 5.

Lincoln documents in archives across the country, notes that his organization once tried using non-academic transcribers, but that the work did not end up being cost effective because archival staff spent so much time and money fixing the volunteers' mistakes. At one point archivists from the project even designed a crowdsourcing platform to transcribe documents, but did not use it because they were "skeptical" about whether it would save any time.⁶⁹ Edward G. Lengel, the editor in chief of a project dedicated to transcribing the works of George Washington, agrees with Stowell and calls crowdsourcing "an unproven concept," pointing out that other project leaders have found their project would have been more cost-effective if their staff had instead devoted their time spent managing volunteers to transcription.⁷⁰ One such example is the Bentham Project, which did eventually become a successful project but had a rocky start. The Bentham Project was founded to allow crowdsourced volunteers to transcribe the works of eighteenth and nineteenth century philosopher Jeremy Bentham. Project staff wrote in the final report that two full-time, temporary staff members were hired to moderate submissions. Staff estimated, however, that these two staff members spent so much time moderating submissions that they would have been able to produce two and a half times as many transcriptions as the volunteers if the two staff members had instead simply devoted their time to transcription. Bentham Project leaders noted, however, that this is a

⁶⁹ Patricia Cohen, "For Bentham and Others, Scholars Enlist Public to Transcribe Papers," *The New York Times*, December 27, 2010, accessed October 20, 2013, <http://www.nytimes.com/2010/12/28/books/28transcribe.html>.

⁷⁰ Marc Parry, "Historians Ask the Public to Help Organize the Past," *The Chronicle of Higher Education*, September 7, 2012, accessed September 8, 2015, <http://chronicle.com/article/Historians-Ask-the-Public-to/134054/>.

pointless argument because they received a grant to fund the project (as well as the temporary staff salaries), and it would be impossible to get a grant if the only goal was transcription without the added digital or participatory aspects.⁷¹

Another popular crowdsourcing venture that many archivists are hesitant to engage in is enlisting the public to create metadata “folksonomies,” or “folk-derived taxonomies.” Folksonomies are metadata created by the public, similar to those that are popular (and are otherwise known as “hashtags,” or just “tags”) on websites such as Flickr, Twitter, and Tumblr.⁷² Folksonomies differ greatly from traditional museum or library taxonomies. Taxonomies are generally hierarchical, controlled, and rigid, and folksonomies are uncentered, informal, and personalized. This informality has the potential to create several problems. Without controlled vocabulary, volunteers may tag images with either one word with multiple meanings, or with several different words with the same meaning, leaving editors to decide whether to include multiple similar, yet slightly different tags. Taggers who are not thinking beyond their own purposes may use abbreviations. Those with different styles may use plurals or singulars, making searching difficult later. Archivists may also run across tags that are spelled incorrectly or are all together wrong. Perhaps because of this, many institutions that have experimented with folksonomies have found that they work best when used alongside a more traditional form of metadata such that the two different forms can complement each other. In this

⁷¹ Tim Causer et al, “Transcription Maximized, Expense Minimized? Crowdsourcing and Editing the Collected Works of Jeremy Bentham,” *Library and Linguistic Computing*, 27 no. 2 (June 2012): 120, 130-131.

⁷² Susan Cairns, “Mutualizing Museum Knowledge: Folksonomies and the Changing Shape of Expertise,” *Curator: The Museum Journal*, 56, no. 1 (2013): 109.

way, professionals who are accustomed to controlled vocabulary can search using traditional terms, and members of the public who are unfamiliar with the rigidity of controlled metadata can search using more intuitive folksonomies.⁷³

While Jascó and Keen both have their fears, examining the history of crowdsourcing and the success of various crowdsourced initiatives shows that the public can indeed provide useful skills and insight to projects. Keen in particular seemed to believe that non-professionals have nothing useful to contribute and that individuals should only focus on what they have been specifically trained and paid to do. However crowdsourced projects have utilized the work of amateurs to achieve great results in many different fields.

CONCLUSION

Crowdsourcing is a new phenomenon whose history and definition are still being written. It had its beginnings in the open source software movement, but has since exploded into countless other industries and endeavors. Authors have been writing about the idea of crowdsourcing for a long time, but Jeff Howe was the first to attach a name to the idea. His article and others present similar ideas and problems associated with crowdsourcing, such as whether this free or cheap labor turns the Internet into a digital sweatshop. Some worry over the results and implications of seeking work from non-professionals.

⁷³ Ibid., 110-111.

Although Keen has his doubts about the public wielding too much control over the Internet and our society, examining the history of crowdsourcing reveals that it is a successful venture for both the solicitor and the volunteer. There is a string of successful projects such as DIY History and Transcribe Bentham to recommend it, and the research of scholars shows that crowdsourcing has been taking place for years, even before it was being called crowdsourcing. As it has evolved, so has its definition. However, it is possible to find a definition that applies to multiple crowdsourcing ventures. Despite the fears of Keen and others, amateurs have proven themselves to be incredibly helpful, as is shown in the following chapter.

CHAPTER IV:

BENEFITS OF CROWDSOURCING IN ACADEMIC INSTITUTIONS

The idea of inviting a seemingly unruly horde of Internet volunteers to tag and transcribe archival records leaves many archivists with understandably mixed feelings. What motivates volunteers to spend their time online working with these historic documents? Can amateurs really be expected to read historical handwriting to produce accurate transcriptions or entrusted with creating useful metadata? Some may worry that their jobs will be lost to online volunteers, or wonder why exactly they paid for years of education when an unknown individual can tag archival materials with whatever words and phrases they choose.

Despite these concerns, crowdsourcing is not a blow to the professionalization of archives, because volunteers are often either performing tasks that are not standard archival work (such as transcription) or are creating complements to archival work (metadata). In this chapter I show that crowdsourcing is a beneficial task by examining the data provided in blog posts, published interviews, and reports from several major crowdsourcing projects, as well as results from conducting interviews with representatives from three institutions managing crowdsourcing projects. To learn what motivates crowdsourcing volunteers, I surveyed volunteers from archival crowdsourcing websites to obtain an approximate demographic base and to determine how their motivations align with academic theories and with the motivations of volunteers in traditional archival settings. Far from being scary hordes, crowdsourcing volunteers are

quite similar to traditional archival volunteers, except that their numbers are much larger. Similar to traditional library and archives volunteers, most crowdsourcing volunteers are motivated by enjoyment of the work and by a strong sense of community and friendship with other volunteers.

Crowdsourcing projects can in fact prove beneficial in both productivity and increased outreach opportunities. Given proper planning, most archives can implement a successful project. This is extremely helpful for today's archives that are faced with staff and budget cuts. Crowdsourcing can help bridge the gap by providing metadata and searchable transcriptions in cases where the collection was not well described. In addition, a strong crowdsourcing project with dedicated volunteers is a wonderful outreach program to help justify the relevance of the institution. This chapter closes with recommendations for archivists considering beginning a crowdsourcing project.

METHODS

To better understand how volunteers on modern crowdsourcing sites relate to traditional volunteers, I conducted an online survey of individuals who volunteer their time on two separate crowdsourcing websites. The survey included ten questions (see appendix) on topics such as age, gender, ethnicity, employment status, household income, and hours per week spent volunteering on a crowdsourcing website. It also included several open-ended questions that asked respondents how they discovered the websites they volunteer with, what their motivations are for volunteering, and whether they volunteer their time on any other similar sites. For legal reasons, the survey was designed to only allow respondents who were over eighteen and who lived within the United States

to complete the entire survey. Incomplete submissions from those who were underaged or outside the country were discarded. I sent copies of this survey to both Project Gutenberg Distributed Proofreaders, a crowdsourcing website on which volunteers proofread scans of historic books that have gone out of copyright in order to create free e-books, and DIY History, a website from the University of Iowa libraries that posts scans of historic letters, diaries, and recipes online and allows volunteers to provide transcriptions. Distributed Proofreaders posted a link to the survey in one of their forums, and DIY History posted the link from their Twitter account. The survey was open from late March to late August, 2014.

To learn more about the pros and cons of using crowdsourcing to complete large projects, I interviewed staff members at three institutions hosting crowdsourcing projects focused on historical and archival projects. Institution A is a non-profit volunteer organization that only receives enough funding to maintain servers to store information for the project. The group uses Optical Character Recognition (OCR) to transcribe historic public domain books, and maintains a large pool of volunteers to edit the OCR transcriptions for errors. These transcriptions are then hosted on a separate website (as opposed to that used for crowdsourcing) in the form of a digital library. Institution B is a small research group that designs educational games. As part of the group's mission, project managers designed a website that displays images provided by libraries, archives, and museums that have chosen to partner with the group. This website allows the public to provide metadata for these images through a digital gaming platform. Volunteers can choose from nine games, some of which are played individually, some with a friend, and

some with a stranger. The information that volunteer gamers input into these games becomes new metadata for the images provided by partner libraries and archives.

Institution C is a medium-sized archives at a large public university. This institution managed a somewhat smaller crowdsourcing project by uploading images of a nearby Civil Rights-era protest to Flickr and asking locals to comment with their memories of the protests as well as the names of those pictured in the photographs. Volunteers instead chose to contact the project manager directly to set up interviews.

“I MAINLY VOLUNTEER BECAUSE IT’S FUN.”¹

The survey received thirty-one total responses. After removing responses of those who were under the age of 18 and those living outside of the United States, twenty four responses remained. In addition, based on the responses to the open-ended questions “How did you find out about the website you volunteer with?” and “Do you volunteer your time on any similar websites?” the majority of respondents seemed to come from Distributed Proofreaders. Few, if any, respondents seemed to be volunteers from DIY History, although it was impossible to be certain as there was no specific question asking respondents to name their volunteer site. This may have led to some homogeneity in responses, especially in responses to motivation, as most survey respondents were experiencing the same atmosphere and digital community. Despite being a small sample size, the survey still offered some useful statistics for comparing crowdsourcing volunteers to historical volunteers. Most individuals who volunteer their time on

¹ Anonymous respondent, “Crowdsourcing: Who Volunteers, and Why?” Internt Survey, April 18, 2014.

crowdsourcing websites do so because they enjoy the work and enjoy the strong sense of community they feel among their fellow volunteers.

Responses to this survey reflected that older adults clearly volunteered more often than their younger counterparts. Only one third of adults reported being under the age of 40, with two even reporting to be over the age of seventy. This conflicts somewhat with the theory set forth in Robert D. Putnam's seminal work *Bowling Alone: The Collapse and Revival of American Community*. Writing in 2000, Putnam argues that since his childhood, America has seen less and less community engagement of all forms, including political engagement (voting, fundraising, and running for office), volunteer activities, and even social activities such as bridge clubs and bowling leagues. He further explains that this disengagement has negative effects on individuals and on society as a whole.² According to Putnam, the biggest factor leading to a decrease in voluntarism and community engagement is generational change, particularly the idea that baby boomers volunteer less than their parents. He argues that the generation of Americans who grew up during the Great Depression and fought in World War II, whom he calls the "long civic generation," learned the importance of community and civic engagement from these disastrous events. Later generations who grew up in relative comfort have never had such an eye-opening example of the need for community association.³

This survey found that the number of hours volunteers were employed did not seem to prevent crowdsourcing volunteers from spending time on crowdsourcing

² Robert D. Putnam, *Bowling Alone: The Collapse and Revival of American Community*, (New York: Simon & Schuster, 2000), 27-28.

³ *Ibid.*, 283-284.

websites. Nearly half of respondents were employed, many full-time, and the amount of time they spent volunteering varied widely (see figure 4). While 21% of respondents said they spent less than five hours a month on the website, nearly 30% reported spending over thirty hours a month volunteering (see figure 5). Even those who were employed full-time still reported devoting long hours to volunteering. While three of the volunteers with full-time employment reported spending less than five hours per month on the site, three volunteers (out of nine responding) reported spending between twenty and thirty hours per month on the website. Two volunteers with full-time work reported spending more than thirty hours per month volunteering. Both respondents who reported working part-time spent less than two hours on the site. Those who volunteered the most were those who were retired; retired individuals reported spending at least ten hours a month on the website and five of the eight surveyed spend thirty hours or more volunteering.⁴ Hours of employment are likely less of an issue for crowdsourcing volunteers than they are for volunteers of organized events because the volunteering takes place online and is time independent. Crowdsourcing is therefore an activity that can be performed at any given time from the volunteer's home, rather than the volunteer needing to maintain set hours and travel to a physical location.

⁴ Ibid.

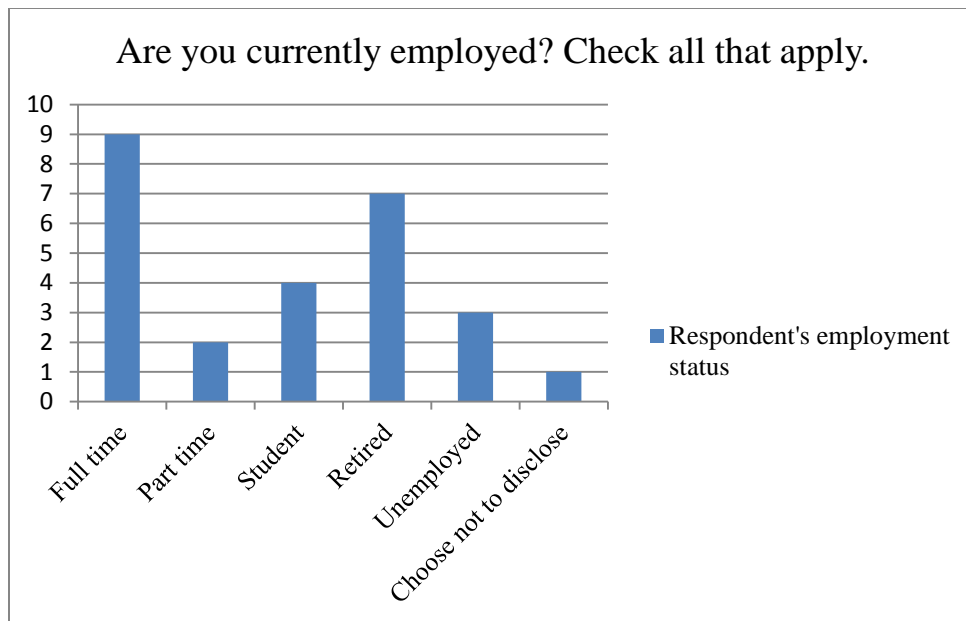


Figure 4. Employment status of survey respondents. Crowdsourcing: Who Volunteers, and Why? Survey by author.

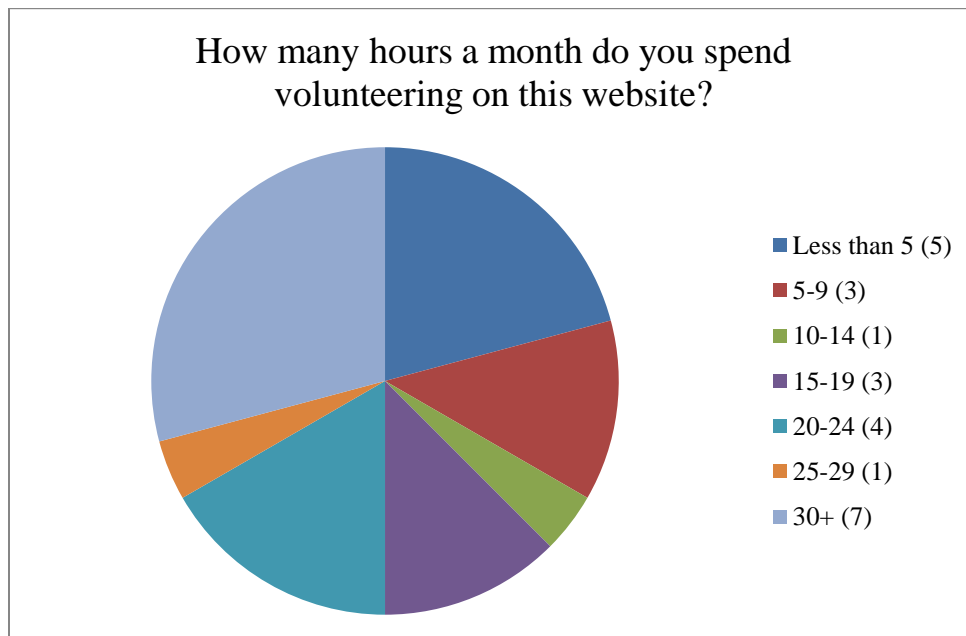


Figure 5. Length of time spent volunteering per month. Crowdsourcing: Who Volunteers, and Why? Survey by author.

The respondents surveyed were widely mixed in their motivation for volunteering. Most survey takers responded with answers that described both altruistic incentives as well as the benefit they received from volunteering (see figure 6). Respondents commonly gave two major motivations for volunteering that could be considered altruistic. The first reason was that they either wanted to preserve history for future generations or make it more accessible for this generation. The second reason was that they felt it was a worthwhile use of their time, or that they had used historical sources from the website they volunteered with in the past and felt the need to give back. Respondents also described several benefits that they received from volunteering. Most volunteered simply because they enjoyed the work involved and many said they took pride in it. Some said they enjoyed learning obscure knowledge from the historical works that they were transcribing and editing. Several respondents also said that they enjoyed the sense of community they felt among the volunteers and the fact that they had been able to “meet,” in a way, new people from around the world.⁵

⁵ Anonymous respondent, “Crowdsourcing: Who Volunteers, and Why?” Internet Survey, April 8, 2014; Anonymous respondent, “Crowdsourcing: Who Volunteers, and Why?” Internet Survey, April 9, 2014.

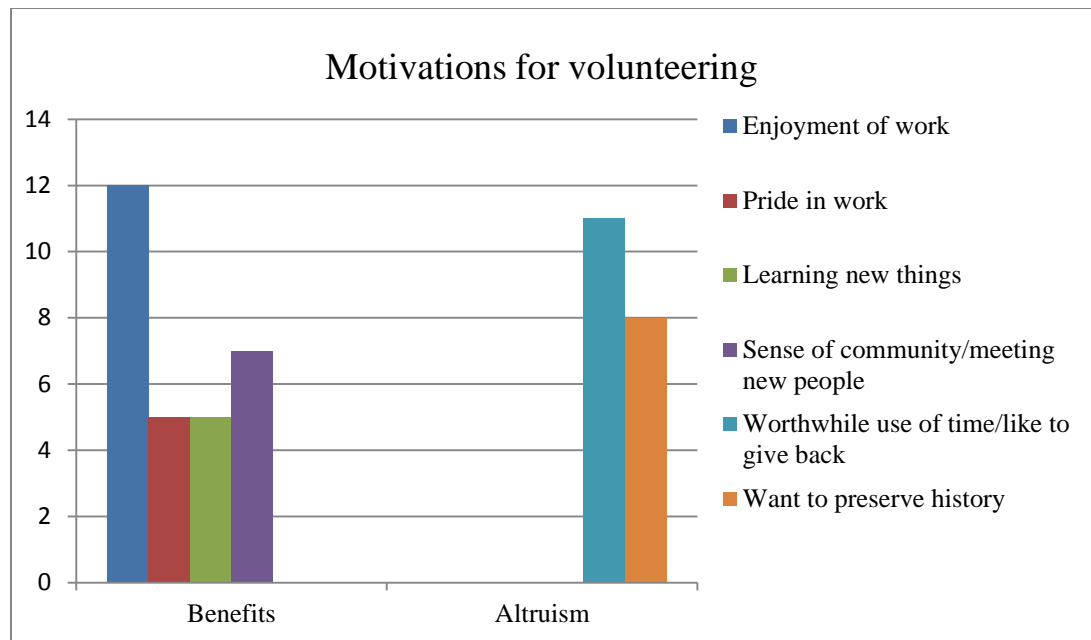


Figure 6. Motivations for volunteering. Crowdsourcing: Who Volunteers, and Why? Survey by author.

According to sociologists Hustinx, Cnaan, and Handy, the motivations acknowledged by crowdsourcing survey takers would likely not be congruent with the expectations of most political scientists and sociologists, who tend to view volunteering as a purely altruistic civic duty. These results, however, are very much in line with economists, who prefer to examine the benefits that one receives for donating one's time to a cause, such as survey respondents who enjoyed learning new information from the works they were transcribing, or those who enjoyed the sense of community among volunteers. These results also align well with the leisure theory of volunteering, which states that altruism is not the principal motivation for volunteering and that it is primarily a leisure activity. One survey taker even responded, "...now it's become a habit and relaxing in an odd sort

of way.”⁶ Another respondent had similar sentiments, saying “I mainly volunteer because it’s fun. It’s a good way to spend time unwinding and taking my mind off of other things.”⁷

DIGITAL VOLUNTEERS AND THE SENSE OF COMMUNITY

As noted previously, some volunteers surveyed discussed the idea of community as central to their volunteering experience. The discovery of a strong sense of community among crowdsourcing volunteers was the most significant result of the survey for two reasons. First, it is entirely contrary to the stereotype that the Internet is populated by lonely and disconnected individuals. Second, it directly ties crowdsourcing volunteers to traditional library and archive volunteers who possess similar motivations. Project Gutenberg Distributed Proofreaders, the organization proofreading digitized versions of historic books, has an extensive web forum that allows its volunteers to discuss the projects they are working on. It also has several forums specifically to foster community, such as “DP [Distributed Proofreaders] Culture and History,” which is described as “A home for discussion of DP as a community, its history and legends, and general DP-oriented chit-chat and entertainment” and a forum called “Everything Else (except DP)”

⁶ Hustinx, Cnaan, and Handy, “Navigating Theories of Volunteering,” 415-419; Stebbins and Graham, *Volunteering as Leisure/leisure as Volunteering*, 7-10; Anonymous respondent, “Crowdsourcing: Who Volunteers, and Why?” Internet Survey, May 30, 2014.

⁷ Anonymous respondent, “Crowdsourcing: Who Volunteers, and Why?” Internet Survey, April 18, 2014.

where members are encouraged to post “recipes, news, general chit-chat, etc. etc.”⁸ One volunteer clearly felt close to the friends she had made by volunteering online: “We have so much fun with teams and we “met” each other even though we were countries apart...I continue to proofread now because it is such a part of my life – it would be like walking away from family.”⁹ Two respondents specifically referred to a sense of community among the volunteers, and, as shown in the chart above, seven noted the importance of community, friendship, or meeting new people. Another volunteer echoed the sentiments of the volunteer above, writing “I have met people in this group from around the country and around the world that I would likely never have come into contact with, and they have enriched my life.”¹⁰ Through these forums, crowdsourcing volunteers have developed strong community bonds. This is contrary to what one might expect, given the perceived anonymous advantages of the Internet. However, these motivations align with traditional volunteers in libraries and archives, many of whom continue their service because of friendships with staff and other volunteers.¹¹

Recent articles have attested that virtual communities do indeed exist and can be evaluated. The article “Sense of Virtual Community: A Follow Up on Its Measurement”

⁸ “Distributed Proofreaders Webforum,” *Project Gutenberg Distributed Proofreaders*, accessed October 14, 2015, <http://www.pgdp.net/phpBB3/index.php>.

⁹ Anonymous respondent, “Crowdsourcing: Who Volunteers, and Why?” Internet Survey, April 8, 2014.

¹⁰ Anonymous respondent, “Crowdsourcing: Who Volunteers, and Why?” Internet Survey, April 9, 2014.

¹¹ Lee Ann Potter and Rebecca Martin, “NARA’s Armies of Volunteers,” in *Prologue* 38 no 4 (Winter 2006), accessed January 13, 2016, <https://www.archives.gov/publications/prologue/2006/winter/volunteers.html>.

attempts to compare virtual communities to “real” communities to determine whether they are similar. The authors first consulted literature to define the phrase “sense of community” as “a feeling that members have of belonging, a feeling that members matter to one another and to the group, and a shared faith that member’s needs will be met through their commitment to the group.”¹² The authors acknowledged that previous studies had found four distinct elements that were vital to developing a sense of community: membership (a sense of safety and belonging), influence (members feel like they have an effect on the community and vice versa), positive reinforcement, and a “shared emotional connection [that] derives from a shared community history, shared events, positive interaction, and identification with the community.”¹³ The authors surveyed members of a German virtual community for the elderly called “Feierabend.de” (“quitting time”) to see how similar or different they were to a “real” community and found several important similarities. The respondents said that the community met their needs, they were able to trust and influence the community; they enjoyed spending time with other community members; and they were hopeful for its future and expected to stay with it for a long time.¹⁴

Authors Anita L. Blanchard and M. Lynne Markus, however, argue that virtual communities are somewhat different from physical communities and have their own

¹² Dagmar Abfalter, Melanie E. Zaglia, and Julia Mueller, “Sense of Virtual Community: A Follow Up on Its Measurement,” *Computers in Human Behavior* 28, no. 2 (March 2012): 400, accessed November 30, 2014, doi:10.1016/j.chb.2011.10.010.

¹³ *Ibid.*, 400-401.

¹⁴ *Ibid.*, 400-402.

experiences. They also use the phrase “sense of virtual community,” but explore the difference between a “sense of community” (SOC) and a “sense of virtual community” (SOVC). The authors developed the idea after realizing that a location or neighborhood was not necessarily a community, but that communities can be a “community of interest” formed around people who share interests, but not neighborhoods, for example collectors or hobbyists. They also noted that there can be a difference between a virtual settlement and a virtual community—a virtual settlement is a website or forum where communication and members exceed a certain amount, but it does not become a community until a level of trust and emotional bonding develops. The authors, therefore, define a sense of virtual community as being “characterized by social processes of 1) exchanging support, 2) creating identities and making identifications, and 3) the production of trust,” saying this is similar to what takes place in offline communities and seems to be what needs to happen to create a sense of virtual community.¹⁵ For their study, the authors examined an online forum to which they assigned the pseudonym “Multiple Sports Newsgroup” (MSN) and sought to determine if members of the site felt a sense of virtual community. They observed posts on the site for seven months and interviewed leaders (who post the most often), participants (who post occasionally), and lurkers (who read posts but never comment). The authors discovered that there were four types of posts on MSN: those that were asking for help, asking and answering questions,

¹⁵ Anita L. Blanchard and M. Lynne Markus, “The Experienced ‘Sense’ of a Virtual Community: Characteristics and Processes,” *Database for Advances in Information Systems* 35, no. 1 (Winter 2004): 66, accessed November 30, 2014, <http://search.proquest.com.ezproxy.mtsu.edu/docview/196638600/624902C095C240F4P/Q/1?accountid=4886>.

etc.; those that were emotion based, such as sharing stories; those that had to do with buying and selling equipment; and those that were just discussions about sports issues. Members that authors interviewed all considered MSN to be a community. The authors noted, however, that the participants' level of participation determined how they interpreted the community. Leaders had met and developed personal friendships with other members and felt the strongest community ties. In contrast, participants and lurkers felt that it was a community in which they were not especially active and in which they had not made personal friendships. Members felt that it was a community because they recognized other members; members created identities for themselves; they received informational and socio-emotional support; and they developed relationships and emotional attachments to other members and "to the community as a whole."¹⁶ Blanchard and Markus concluded that the MSN is a SOVC because it contains the community-like aspects of exchanging support, creating identities, and producing trust. They acknowledge, however, that the SOVC differs from the SOC in that the factor of influence disappears and that there is more individualism online than offline—members create identities that align with the group, rather than simply identifying with the group. The authors explain that this may be an attempt to combat the anonymity of the Internet and not get lost in the crowd.¹⁷

Are the members of DIY History and Distributed Proofreaders engaged in true virtual communities? Abfalter, Zaglia, and Mueller's definition centers on the types of

¹⁶ Blanchard and Markus, "The Experienced 'Sense' of a Virtual Community", 66, 71-73.

¹⁷ Ibid., 73-75.

communication and interaction between members. In order to meet this definition of a virtual community, members should feel a sense of belonging and influence on each other and the community, and the community should have a shared history.¹⁸ DIY History likely does not meet this definition as there are no forums for members to hold discussions, and the only way for members to converse is in the comments section on individual pages to be transcribed. Very little interaction therefore takes place and most discussion consists of interested comments on manuscripts (for example: “This apple pudding sounds delicious!”)¹⁹ Distributed Proofreaders, however, meets this definition very well. Members surveyed have said that those they met through the community “have enriched my life” and that leaving the group would be “like walking away from family.”²⁰ As noted earlier, the website also has a forum specifically reserved as “a home for discussion of DP as a community, its history and legends.”²¹ Maintaining some sort of forums or direct messaging system where members can meet to make friends and feel welcome discussing topics other than the manuscripts and tasks at hand is vital to creating a strong sense of community. It is also interesting, and perhaps important, that

¹⁸ Abfalter, Zaglia, and Mueller, “Sense of Virtual Community,” 400-401.

¹⁹ Andrew Parker, comment on Marie Carnegie and Susan Gillespie cookbook, pg 93, *DIY History*, December 19, 2014, accessed March 7, 2016, <http://diyhistory.lib.uiowa.edu/transcribe/2982/63723>.

²⁰ Anonymous respondent, “Crowdsourcing: Who Volunteers, and Why?” Internet Survey, April 9, 2014; Anonymous respondent, “Crowdsourcing: Who Volunteers, and Why?” Internet Survey, April 8, 2014.

²¹ “Distributed Proofreaders Webforum,” *Project Gutenberg Distributed Proofreaders*, accessed October 14, 2015, <http://www.pgdp.net/phpBB3/index.php>.

respondents were quick to call the group a community without any prompt or mention of the word in the survey that they were given.

This thesis demonstrates that crowdsourcing volunteers are similar to traditional volunteers in libraries and archives in their enjoyment either of the community aspect of volunteering or of simply working with the records themselves. As discussed in chapter one, many who volunteer in person at a library or archive say that their friendships with staff and other volunteers is what keeps them coming back to the activity, while others are simply “history buffs” who enjoy the opportunity to work with historic records they would not otherwise have access to.²² Similarly, many crowdsourcing volunteers surveyed were motivated to volunteer because they enjoyed the work or because they enjoyed making friends through forums on crowdsourcing websites.²³ Archivists should be aware of these motivations as a way to recruit and maintain crowdsourcing volunteers. New projects could be announced across various web platforms and forums devoted to history in an attempt to attract those who may be interested in working with historical records. In addition, building discussion forums into new projects can help members build a sense of community that may retain volunteers for the lifespan of the project.

²² Lee Ann Potter and Rebecca Martin, “NARA’s Armies of Volunteers,”; Stebbins, introduction, 3-4, 25-26.

²³ Anonymous respondent, “Crowdsourcing: Who Volunteers, and Why?” Internet Survey, April 9, 2014.

“THAT’S A KEY WORD, *ENGAGE*. THAT’S WHY WE EXIST.”²⁴

Crowdsourcing projects can reap many benefits for both the institution and the volunteers. Diverse crowds tagging materials can create higher quality metadata, at least in terms of making collections more searchable to the general public. If used well, crowdsourcing can also improve archives’ efficiency in completing tasks. Perhaps most importantly, the community engagement opportunities it creates and the pride of ownership that volunteers feel in materials they helped create help keep quality crowdsourcing projects going strong.²⁵

Although crowdsourced folksonomies, or tagging, take some of the work of archivists out of the hands of professionals, they can also be extremely helpful. Digitizing collections is a wonderful idea for an institution attempting to make their materials more accessible. However, this step does not necessarily ensure these materials will be accessible if the metadata is full of professional terminology to the point that the public cannot navigate a simple search.²⁶ For example, members of the public who do not remember the name of the artwork or artist they are searching for likely will not also remember when the artwork was created or that the museum metadata included the words “enameled,” “gilt” or “metalwork.” When asked if the tags created by their volunteers were of good quality, a staff member at Institution B, a small research group that built a website allowing volunteers to play games that would supply metadata for library,

²⁴ Staff member at Institution C, Skype interview with the author, October 1, 2014.

²⁵ Ellis, “A History of Collaboration,” 4.

²⁶ Cairns, “Mutualizing Museum Knowledge, 110

archival, and museum collections turned the question back to me. “With tags, what constitutes good quality?” they asked, implying that any folksonomic data that will improve searchability is good data.²⁷ Allowing crowdsourced volunteers to add folksonomic metadata helps non-professionals who do not understand controlled vocabularies search digital collections with greater speed and accuracy. They can also serve to help the public build a narrative of sorts because they “introduce previously unconsidered perspectives by recording an individual user’s personal response to the object.”²⁸ It is also important to note that these folksonomies are not replacing professional metadata, but rather complementing it. In this way there remains a predictable, controlled vocabulary for archives professionals to use when searching and a simpler vocabulary for the public to use when searching. With the rise in community archives, open access to records, and now crowdsourcing, archival professionals have repeatedly worried over the loss of professionalism in archives. Rather than a loss of professionalization, crowdsourcing represents simply a new option for outreach and collaboration.

Several examples exist of successful crowdsourced tagging projects. The *steve.museum* project was a collaboration between staff members at several different large institutions including the Metropolitan Museum of Art, the Guggenheim Museum, The Cleveland Museum of Art, and the San Francisco Museum of Modern Art, that allowed social tagging of collections at art museums. Museum professionals pioneered

²⁷ Staff member at Institution B, telephone interview with the author, November 6, 2014.

²⁸ Cairns, “Mutualizing Museum Knowledge, 110, 115.

the project after realizing that members of the public were experiencing extreme difficulty with searching digital collections.²⁹ The project was tested for two years between 2006 and 2008, and researchers found that 86% of the tags created were not replicated in museum metadata. In addition, “museum staff felt that 88% of taggers’ terms would be useful for searching.”³⁰ A more recent example is the website “The Metadata Games,” which hosts images for libraries and archives and allows users to suggest tags through gaming interfaces in hopes of attracting those who would not normally volunteer on crowdsourcing sites. In their report for their initial testing period, the authors noted that they looked to projects like the Library of Congress’ (LOC) Flickr tagging project as a potential model. LOC began placing photographs on the photo sharing website Flickr in January 2008 with mixed hopes of increasing public awareness of LOC holdings, experimenting with social tagging, and potentially gaining useful information from the tags and comments.³¹ They found that the LOC’s project was very successful both in terms of numbers and content. For example, for several photos, volunteers had added the tag “Rosie the Riveter” where the LOC’s only tags were “Women—employment” and “World War, 1939-1945.” During the pilot test for their

²⁹ Ibid., 110; Bruce Wyman et al, “Steve.museum: An Ongoing Experiment in Social Tagging, Folksonomy, and Museums,” in J. Trant and D. Bearman (eds.), *Museums and the Web 2006: Proceedings*, Toronto: Archives & Museum Informatics, published March 1, 2006, accessed January 22, 2016, <http://www.archimuse.com/mw2006/papers/wyman/wyman.html>.

³⁰ Cairns, “Mutualizing Museum Knowledge,” 110.

³¹ Michelle Springer et al, “For the Common Good: The Library of Congress Flickr Pilot,” (Library of Congress, October 30, 2008), accessed January 23, 2016, http://www.loc.gov/rr/print/flickr_report_final.pdf.

own site, they found that some photographs received similar tags to what they had been previously assigned in metadata, but others gathered tags that could be helpful, such as the specific name of a plant growing in a photograph rather than just the word “shrub” or a photograph with the word “family,” when the metadata had not otherwise specified this information. In all, the researchers found that by using games as motivation, they received more tags as compared to the LOC. The LOC received an average of seventeen to eighteen tags per image and .006 tags per user, but the Metadata games received thirty-two to thirty-three tags per image and .84 tags per person.³²

Crowdsourced volunteers can complete work quickly, efficiently, and cost-effectively, if managed well. The Civil War Diaries Transcription Project (later known as DIY History) at the University of Iowa was created because the staff was interested in digitized documents being transcribed so that they could be more easily searched, but they wanted to save money because they were not receiving any extra funding to finish the project. Staff members noted that it took longer to check volunteer work than it would have taken to check the work of professional transcribers, but that it did not take extremely long and it cost much less. In an interview for an LOC blog post, Nicole Saylor, the head of Digital Library Services at the University of Iowa, said that she felt the project had been very successful and had attracted loyal volunteers, explaining that one volunteer had transcribed more than five hundred pages. The assistant head of special collections for the library agreed, saying that they had “come to recognize some ‘power

³² Mary Flanagan and Peter Carini, “How Games Can Help Us Access and Understand Archival Images,” *American Archivist* 75 no. 2 (October 2012): 521, 518, 529-532.

users' who transcribe in great quantity with high accuracy."³³ Saylor also mentioned that one of their volunteers began offering corrections to earlier work: "evidence that you should never underestimate the crowd."³⁴

In contrast to how others might perceive it, staff members at the Bentham Project view their crowdsourcing endeavor as a success. Prior to beginning crowdsourcing, it took the project fifty years to transcribe twenty-seven of what they estimate will be seventy volumes of Bentham's writings. For the crowdsourcing project, staff scanned nearly forty thousand pages of the remaining manuscripts and put them online. During the first six months, 1,207 registered users transcribed 1,009 manuscripts, 56% of which staff members determined were complete. Interestingly, the amount of work completed by volunteers was not at all uniform—only 21% of registered users actually completed any transcriptions, and two-thirds of those users only completed one page. One-quarter of users completed between two and five pages, fifteen users transcribed between six and thirty pages, six users completed between sixty three and eighty two pages each, and one volunteer completed 280 pages.³⁵ Although skeptics have pointed to calculations that staff members could have completed more work if they had transcribed full time, those at

³³ Jie Jenny Zou, "Civil War Project Shows Pros and Cons of Crowdsourcing," *The Chronicle of Higher Education*, June 14, 2011, accessed May 24, 2013, <http://chronicle.com/blogs/wiredcampus/civil-war-project-shows-pros-and-cons-of-crowdsourcing>; Trevor Owens and Bill LeFurgy, "Crowdsourcing the Civil War: Insights Interview With Nicole Saylor," *The Signal: Digital Preservation*, December 6, 2011, accessed May 24, 2014, <http://blogs.loc.gov/digitalpreservation/2011/12/crowdsourcing-the-civil-war-insights-interview-with-nicole-saylor/>; Zou, "Civil War Project."

³⁴ Owens and LeFurgy, "Crowdsourcing the Civil War."

³⁵ Cohen, "For Bentham and Others."; Causer et al, "Transcription Maximized, Expense Minimized?" 125-126.

the Bentham Project argue that their results from the six-month pilot study are “quite remarkable,” considering the level of difficulty involved in reading the handwriting and marginalia in the documents plus the sheer length of each documents (Each page averaged 250-750 words, but some were up to two thousand words long.)³⁶ The Bentham Project currently posts weekly reports to their blog. As of this writing, the most recent update was posted on September 4, 2015 and announced that 15,176 out of 35,002 pages had been transcribed, a total of 43%. Ninety-three percent of those pages had been approved as being complete and correct by staff members. The post also included a chart indicating the number of pages transcribed in every box in the collection. The collection contains 176 boxes—the box with the fewest number of pages transcribed had not been started. Six of the boxes, however had been completely transcribed.³⁷

The employees interviewed for this thesis have also found success in their crowdsourcing projects. Institution A has 127,066 volunteers currently registered with their project and has completed transcribing and editing 28,261 complete books.³⁸ Institution B registered 240 users over nearly one year, however most of their metadata came from unregistered users. Approximately fifteen thousand games had been played during the year and they have collected twenty five thousand distinct tags on thirteen

³⁶ Causer et al, “Transcription Maximized, Expense Minimized?” 127.

³⁷ Louise Seaward, “Progress update, 30 Jan to 5 Feb”, *Transcribe Bentham*, February 5, 2016, accessed February 11, 2016, <http://blogs.ucl.ac.uk/transcribe-bentham/2016/02/05/progress-update-30-jan-to-5-feb/>.

³⁸ Manager at Institution A, email interview with the author, August 2014.

thousand items.³⁹ Institution C has found success, but not in the way they initially expected. They posted images on Flickr and waited for comments, but instead those who had been involved in the Civil Rights protests called and asked to speak to them. Staff ended up interviewing approximately twenty five former protestors and their families, often spending two hours in each interview, and identified about sixty people from the photographs as well as learning more about the protest and its impacts on the community.⁴⁰

The ability for large groups of volunteers to achieve quick results is especially important in today's world of understaffed and underfunded archives facing huge backlogs. In 2005, archivists Mark A. Greene and Dennis Meissner attempted to solve the problem of enormous archival backlogs by outlining a method for faster processing which came to be called "More Product, Less Process" (MPLP). One result of their method was a decreased level of arrangement and description. Archivists had mixed responses to this method. Some loved the idea; some argued it was not actually a new idea; and some worried the idea would go too far and poorly described records would make it difficult later to help researchers during reference requests.⁴¹ In the way it is currently being used, crowdsourcing can help solve the description problem by increasing

³⁹ Staff member at Institution B, telephone interview with the author, November 6, 2014.

⁴⁰ Staff member at Institution C, Skype interview with the author, October 1, 2014.

⁴¹ Mark A. Greene and Dennis Meissner, "More Product, Less Process: Revamping Traditional Archival Processing," *The American Archivist* 68, no. 2 (2005): 208–63, accessed February 12, 2016, <http://www.jstor.org/stable/40294290>; Carl Van Ness, "Much Ado about Paper Clips: 'More Product, Less Process' and the Modern Manuscript Repository," *The American Archivist* 73, no. 1 (2010): 129–45, accessed February 12, 2016, <http://www.jstor.org/stable/27802718>.

access to documents that were processed using MPLP and that, as a result, might have poor finding aids. Transcription of scanned documents allows the full text to be searchable for subjects, names, and places that may not have been included in finding aids. There is initially a large amount of work involved in scanning documents. Subsequently, however, reference work, especially for genealogists looking for names of specific ancestors, becomes much easier.

Perhaps the most important, and often overlooked, outcome of crowdsourcing ventures is increased visibility of archives and engagement between users and staff. Outreach and engagement with a larger segment of public, rather than just professional historians and scholars, has come front and center in archival literature in recent years, especially as budgets shrink and institutions find themselves looking to justify their existence. Well-managed crowdsourcing programs can make for incredible outreach opportunities. At the end of their six-month trial period, the staff at the Bentham Project said that their project should be viewed not for the number of pages transcribed but instead for the way it publicized archives and crowdsourcing. Sharon M. Leon, director of public projects at the Center for History and New Media at George Mason University, argues that crowdsourcing is a good outreach tool that illustrates the importance of history to the public and improves public involvement. It also creates different kinds of work for archivists, rather than taking work away from them, and she argues that this should be well marketed in projects so the public does not worry.⁴² Institution C found outreach to the public to be “a huge benefit” of their crowdsourcing project. One of their

⁴² Parry, “Historians Ask the Public.”; Zou, “Civil War Project.”

motivations for beginning the project was to establish goodwill between the community and the institution, and they more than succeeded. “It’s a wonderful thing to have the public *engage* with your collection. And that’s a key word, *engage*,” the project manager said. “That’s the reason we exist, not just to preserve, but to engage.”⁴³ The manager was also especially adamant when stating:

Crowdsourcing represents libraries, archives, et cetera, putting a toe into the new world of the web. The web is about a conversation, not a broadcast. Institutions have historically been most comfortable with broadcast and this is a way for us to learn a new mode of interaction with the public. I hope that we take these lessons and become new and stronger institutions.⁴⁴

Crowdsourcing is successful as an outreach program because interested volunteers can not only view, but also interact with primary source historical documents. Although the volunteers with Institution C chose to visit the archives to participate, this is not a requirement with a traditional crowdsourcing platform, and volunteers can therefore interact with the materials from thousands of miles away at a time and location of their choosing. This allows those who may not be able to travel to visit an institution in person the ability to view historical materials first hand. Crowdsourcing is different from simply placing a series of photographs or letters into a digital exhibit. Rather than being asked to passively look at a few letters or photographs, volunteers are encouraged to perform a task in which they actively engage and analyze a document. Nicole Saylor, who helped found the website DIY History, discovered that many volunteers “became invested in the story” in the diaries they were transcribing, and one volunteer sent a message to staff

⁴³ Staff member at Institution C, Skype interview with the author, October 1, 2014.

⁴⁴ Staff member at Institution C, Skype interview with the author, October 1, 2014.

saying “This is one of the COOLEST and by far most historically significant things I have seen since I first saw a dinosaur fossil and realized how big they are.”⁴⁵

CONCLUSION

Surveyed volunteers on two different websites were found to have motivations for volunteering and ideas of community that aligned with academic theories, as well as with traditional archives volunteers. The predictions of economists align best with the motivations of crowdsourcing volunteers, in that economists argue volunteers are motivated by that sort of “warm fuzzy feeling” one gets from doing good deeds. However crowdsourcing volunteers also confirm the theory that volunteering is an enjoyable activity that many participate in for entertainment. Survey results show that volunteers participate on crowdsourcing websites because they enjoy the work; they enjoy being able to do something worthwhile; and they enjoy the sense of community and friendship they feel on crowdsourcing websites. Recently, authors have argued that these new technologies have formed a new sort of community around them, a virtual community, that is different but no less rewarding than a community which meets face-to-face. Survey respondents agreed and interpreted their own group as a sort of community in which members meet to discuss and proofread historic books.

Many scholars and archivists hold deep concerns about whether well-intentioned volunteers can be entrusted with tasks such as providing metadata and transcribing difficult historical handwriting. Research and interviews with those managing successful

⁴⁵ Owens and LeFurgy, “Crowdsourcing the Civil War.”

crowdsourcing projects confirm that the pros outweigh the cons. Although crowdsourcing often requires extensive staff involvement, it can yield quick results and allow archives to receive funding for projects that would not otherwise be eligible for funds. Best of all, crowdsourcing is an excellent way to increase public awareness and use of digital collections. Not only will larger groups become aware of the institution's digital holdings, but also the digital holdings themselves often become more easily searchable as volunteers add folksonomic metadata. While volunteers are closely examining the historic materials to determine what to transcribe or what tags to add, they interact with that record at a level that previously the public was rarely able to, and often develop a new appreciation for the historical record in the process. Archivists should not fear or worry over crowdsourcing projects, but rather approach crowdsourcing with enthusiasm and careful planning to ensure success in their project.

BIBLIOGRAPHY

PRIMARY SOURCES

- “About FamilySearch.” FamilySearch.org. Accessed November 3, 2013.
<https://familysearch.org/about>.
- “About Madison.” The New York Times Madison. Accessed March 14, 2016.
<http://madison.nytimes.com/contribute>.
- “About the Ancestry.com World Archives Project.” Ancestry.com. Accessed November 3, 2013. <http://landing.ancestry.com/wap/learnmore.aspx>.
- “About the Polar Bear Expedition Digital Collections.” Polar Bear Expedition Digital Collections. Accessed February 12, 2016.
<http://quod.lib.umich.edu/p/polaread/about.html>.
- “ALA Standards & Guidelines.” *American Library Association*, March 26, 2015. Accessed February 19, 2016.
<http://www.ala.org/tools/guidelines/standardsguidelines>.
- Ancestry.com. “Ancestry.com Announces the World Archives Project.” *Ancestry.com Blog*, September 4, 2008. Accessed November 3, 2013.
<http://blogs.ancestry.com/ancestry/2008/09/04/ancestrycom-announces-the-world-archives-project/>.
- “ArchivesSpace Home.” ArchivesSpace. Accessed February 20, 2016.
<http://www.archivesspace.org/>.
- “Archivists’ Toolkit.” Archivist’s Toolkit. Accessed February 20, 2016.
<http://www.archiviststoolkit.org/>.
- “Archon: The Simple Archival Information System.” Archon. Accessed February 20, 2016. <http://www.archon.org/>.
- “Behind-the-Scenes Volunteer Program.” Smithsonian. Accessed March 5, 2016.
<http://www.si.edu/Volunteer/Behind-the-Scenes-Volunteer>.
- Carroll, Sean. “Site of the Week: Wikipedia.” *PC Magazine.com*, June 6, 2003. Accessed November 1, 2013.
<http://www.lexisnexis.com.ezproxy.mtsu.edu/hottopics/lnacademic/>.
- Causser, Tim, Justin Tonra, and Valerie Wallace. “Transcription Maximized; Expense Minimized? Crowdsourcing and Editing The Collected Works of Jeremy Bentham.” *Literary & Linguistic Computing* 27, no. 2 (June 2012): 119–37. Accessed May 22, 2014.
<https://ezproxy.mtsu.edu:3443/login?url=http://search.ebscohost.com.ezproxy.mtsu.edu>

u.edu/login.aspx?direct=true&db=edswah&AN=000304199900001&site=eds-live&scope=site.

Cohen, Patricia. "For Bentham and Others, Scholars Enlist Public to Transcribe Papers." *The New York Times*, December 27, 2010, sec. Books. Accessed October 20, 2013. <http://www.nytimes.com/2010/12/28/books/28transcribe.html>.

Crowdsourcing: Who Volunteers, and Why? Interview by Kayla Utendorf. Internet Survey, August 2014.

"Distributed Proofreaders Web Forum." *Project Gutenberg Distributed Proofreaders*. Accessed October 15, 2015. <http://www.pgdp.net/phpBB3/index.php>.

Douglas, Nick. "Job Market News: That's Not Slave Labor, That's Crowdsourcing!" *Valleywag [Gawker]*, May 25, 2006. Accessed February 6, 2016. http://www.lexisnexis.com.ezproxy.mtsu.edu/lnacui2api/results/docview/docview.do?start=2&sort=BOOLEAN&format=GNBFI&risb=21_T18420586719.

———. "Wagged, Sagged, Body-Bagged: Things We've Decided Are Dead." Media Blog. *Valleywag [Gawker]*, May 25, 2006. Accessed October 19, 2013. <http://gawker.com/176180/wagged-sagged-body+bagged-things-weve-decided-are-dead>.

Flanagan, Mary. "Who We Are." Crowd Consortium. Accessed March 14, 2016. <http://www.crowdconsortium.org/who-we-are-2/>.

"Frequently Asked Questions." Old Weather. Accessed September 11, 2015. <http://www.oldweather.org/faq>.

"Guidelines for Using Volunteers in Libraries." *American Libraries* 2, no. 4 (1971): 407–8. Accessed February 19, 2016. <http://www.jstor.org/stable/25618274>.

Howard, Jennifer. "Breaking Down Menus Digitally, Dish by Dish." *Chronicle of Higher Education* 58, no. 35 (May 4, 2012): B18–19. Accessed May 24, 2013. <http://ehis.ebscohost.com.ezproxy.mtsu.edu/ehost/detail?vid=8&sid=a196ef29-ef86-458f-8c82-8bbcbf9761f4%40sessionmgr114&hid=102&bdata=JnNpdGU9ZWwhvc3QtbGl2ZSZzY29wZT1zaXRl#db=trh&AN=75230240>.

Howe, Jeff. "5 Rules of the New Labor Pool." *Wired* 14, no. 6 (June 2006). Accessed October 20 2013. <http://www.wired.com.ezproxy.mtsu.edu/wired/archive/14.06/labor.html>.

———. "Birth of a Meme." Blog. *Crowdsourcing.com*, May 27, 2006. Accessed October 20 2013. http://www.crowdsourcing.com/cs/2006/05/birth_of_a_meme.html.

———. "Crowdsourcing." Blog. *Crowdsourcing.com*. Accessed October 20, 2013. <http://www.crowdsourcing.com/cs/>.

- . “Crowdsourcing: A Definition.” Blog. *Crowdsourcing.com*, June 2, 2006. Accessed October 20, 2013. http://www.crowdsourcing.com/cs/2006/06/crowdsourcing_a.html.
- . “The Rise of Crowdsourcing.” *Wired* 14, no. 6 (June 2006). Accessed October 19, 2013. <http://www.wired.com.ezproxy.mtsu.edu/wired/archive/14.06/crowds.html>.
- “Indexing Overview.” FamilySearch.org. Accessed February 2, 2016. <https://familysearch.org/indexing/>.
- “InnoCentive.” Innocentive. Accessed March 7, 2016. <http://www.innocentive.com/>.
- Interviews with Institutions Hosting Crowdsourcing Projects. Interview by Kayla Utendorf, Fall 2014.
- Jascó, Péter. “Péter’s Picks & Pans.” *Online Magazine* 26, no. 2 (April 2002): 79–82. Accessed November 1, 2013. <https://ezproxy.mtsu.edu:3443/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=llf&AN=502875163&site=eds-live&scope=site>.
- Levack, Kinley. “If Two Heads Are Better than One, Try 7,000 with Wikipedia.” *EContent* 26, no. 4 (April 2003): 12–13. Accessed November 1, 2013. <https://ezproxy.mtsu.edu:3443/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=llf&AN=501020266&site=eds-live&scope=site>.
- “Main Page.” Wikipedia, the Free Encyclopedia. Accessed February 2, 2016. http://en.wikipedia.org/wiki/Main_Page.
- Naruto-Moya. “Welcome! Help Build the Indigenous Digital Archive.” Indigenous Digital Archive, September 17, 2014. Accessed March 13, 2016. <http://blog.indigenousdigitalarchive.org/post/2014/09/17/first>.
- NYPL Labs. “What’s on the Menu? About.” New York Public Library. Accessed October 21, 2013. <http://menus.nypl.org/about>.
- “Omeka.” Omeka. Accessed February 20, 2016. <http://omeka.org/>.
- Owens, Trevor, and Bill LeFurgy. “Crowdsourcing the Civil War: Insights Interview with Nicole Saylor.” *The Signal: Digital Preservation*, December 6, 2011. Accessed May 24 2013. <http://blogs.loc.gov/digitalpreservation/2011/12/crowdsourcing-the-civil-war-insights-interview-with-nicole-saylor/>.
- Parry, Marc. “Historians Ask the Public to Help Organize the Past.” *Chronicle of Higher Education* 59, no. 2 (September 7, 2012): 17–17. Accessed September 8 2015. <http://chronicle.com/article/Historians-Ask-the-Public-to/134054/>.
- “Scripto.” Scripto. Accessed February 20, 2016. <http://scripto.org/>.

- Seaward, Louise. "Progress Update, 30 Jan to 5 Feb." *Transcribe Bentham*, February 5, 2016. Accessed February 11, 2016. <http://blogs.ucl.ac.uk/transcribe-bentham/2016/02/05/progress-update-30-jan-to-5-feb/>.
- Society of American Archivists. "Best Practices for Volunteers in Archives." *Society of American Archivists*, August 2014. Accessed January 14, 2016. <http://www2.archivists.org/standards/best-practices-for-volunteers-in-archives>.
- Springer, Michelle, Beth Dulabahn, Phil Michel, Barbara Natanson, David Reser, David Woodward, and Helena Zinkham. "For the Common Good: The Library of Congress Flickr Pilot Project." Library of Congress, October 30, 2008. Accessed January 23 2016. http://www.loc.gov/rr/print/flickr_report_final.pdf.
- "The Open Source Definition (Annotated)." Open Source Initiative. Accessed November 4, 2013. <http://opensource.org/osd-annotated>.
- The University of Iowa Libraries. "DIY History." DIY History: The University of Iowa Libraries, 2012. Accessed December 9, 2013. <http://diyhistory.lib.uiowa.edu/>.
- "Transcribe Bentham: Transcription Desk." University College London, December 6, 2013. Accessed December 8 2013. http://www.transcribe-bentham.da.ulcc.ac.uk/td/Transcribe_Bentham.
- "Volunteer." Jefferson County Public Library, 2016. Accessed March 5 2016. <http://jeffcolibrary.org/volunteer>.
- "Volunteer." Des Moines Public Library. Accessed March 5, 2016. <http://dmpl.org/volunteer>.
- "Volunteer Opportunities." San Diego County Library. Accessed March 5, 2016. <http://www.sdcl.org/volunteer.html>.
- "Volunteer Programs." Maryland State Archives, January 27, 2016. <http://msa.maryland.gov/msa/refserv/html/volunprogram.html>.
- Watson, Richard. "Tech@Work: A Problem Shared Is a Problem Solved." Business Blog. *The Edge Singapore*, January 24, 2005. Accessed October 19, 2013. <http://www.lexisnexis.com.ezproxy.mtsu.edu/hottopics/lnacademic/>.
- Zou, Jie Jenny. "Civil War Project Shows Pros and Cons of Crowdsourcing." *The Chronicle of Higher Education*, June 14, 2011. Accessed May 24 2014. <http://chronicle.com/blogs/wiredcampus/civil-war-project-shows-pros-and-cons-of-crowdsourcing>.

SECONDARY SOURCES

- Abfalter, Dagmar, Melanie E. Zaglia, and Julia Mueller. "Sense of Virtual Community: A Follow Up on Its Measurement." *Computers in Human Behavior* 28, no. 2 (March 2012): 400–404. Accessed November 30, 2014. doi:10.1016/j.chb.2011.10.010.
- Adrienne C. Thomas. "Our Wonderful Volunteers." *Prologue* 41, no. 3 (Fall 2009): 4. Accessed January 12 2016.
<https://www.archives.gov/publications/prologue/2009/fall/archivist.html>.
- Anderson, Benedict R. *Imagined Communities: Reflections on the Origin and Spread of Nationalism*. Rev. ed. London ; New York: Verso, 2006.
- Bailey, Beth L. *America's Army: Making the All-Volunteer Force*. Cambridge: Harvard University Press, 2009.
- Baytiyeh, Hoda, and Jay Pfaffman. "Volunteers in Wikipedia: Why the Community Matters." *Journal of Educational Technology & Society* 13, no. 2 (July 2010): 128–40. Accessed February 2, 2016.
<https://ezproxy.mtsu.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=aph&AN=52045409&site=eds-live&scope=site>.
- Benkler, Yochai. "Coase's Penguin, Or, Linux and 'The Nature of the Firm.'" *The Yale Law Journal* 112, no. 3 (December 1, 2002): 369–446. Accessed October 20, 2013. doi:10.2307/1562247.
- Blanchard, Anita L., and M. Lynne Markus. "The Experienced 'Sense' of a Virtual Community: Characteristics and Processes." *Database for Advances in Information Systems* 35, no. 1 (Winter 2004): 65–79. Accessed November 30, 2014.
<http://search.proquest.com.ezproxy.mtsu.edu/docview/196638600/624902C095C240F4PQ/1?accountid=4886>.
- Bretthauer, David. "Open Source Software: A History." *Information Technology & Libraries* 21, no. 1 (March 2002): 3–10. Accessed December 9, 2013.
<https://ezproxy.mtsu.edu:3443/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=aph&AN=6607909&site=eds-live&scope=site>.
- Cairns, Susan. "Mutualizing Museum Knowledge: Folksonomies and the Changing Shape of Expertise." *Curator: The Museum Journal* 56, no. 1 (2013): 107–19. Accessed January 31, 2015. doi:10.1111/cura.12011.
- Cash, Floris Loretta Barnett. *African American Women and Social Action: The Clubwomen and Volunteerism from Jim Crow to the New Deal, 1896-1936*. Contributions in Women's Studies: No. 188. Westport, CT: Greenwood Press, 2001.

- Cobbs Hoffman, Elizabeth. *All You Need Is Love: The Peace Corps and the Spirit of the 1960s*. Cambridge: Harvard University Press, 2000.
- “Crowdsourcing, N.” *OED Online*. Oxford University Press. Accessed February 19, 2016. <http://www.oed.com.ezproxy.mtsu.edu/view/Entry/376403#eid288590739>.
- DiBona, Chris, Sam Ockman, and Mark Stone, eds. *Open Sources: Voices from the Open Source Revolution*. Sebastopol, CA: O’Reilly Media, Inc., 1999.
- DiBona, Chris, Mark Stone, and Danese Cooper. *Open Sources 2.0: The Continuing Evolution*. 1 edition. Sebastopol, CA: O’Reilly Media, 2005.
- Doraiswamy, Uma. “Tips for Library and Information Science Students Seeking Employment and Entering the Workforce.” *Collaborative Librarianship* 3, no. 3 (July 2011): 176–79. Accessed February 19, 2016. <https://ezproxy.mtsu.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=ofm&AN=67046968&site=eds-live&scope=site>.
- Dvorak, John C. “Upstarts Attack Microsoft Slackers.” *PC Magazine*, 2002. Accessed December 9, 2013. <https://ezproxy.mtsu.edu:3443/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=edsgao&AN=edsgcl.133520159&site=eds-live&scope=site>.
- Ellis, Sally. “A History of Collaboration, a Future in Crowdsourcing: Positive Impacts of Cooperation on British Librarianship.” *Libri: International Journal of Libraries & Information Services* 64, no. 1 (March 2014): 1–10. Accessed April 7, 2014. doi:10.1515/libri-2014-0001.
- Estellés-Arolas, Enrique, and Fernando González-Ladrón-de-Guevara. “Towards an Integrated Crowdsourcing Definition.” *Journal of Information Science* 38, no. 2 (April 2012): 189–200. Accessed November 3, 2013. doi:10.1177/0165551512437638.
- Ewing, Susan E. “Using Volunteers for Special-Project Staffing at the National Air and Space Museum Archives.” *The American Archivist* 54, no. 2 (1991): 176–83. Accessed January 12, 2016. <http://www.jstor.org/stable/40293550>.
- Flagg, Rachel. “Basics of Survey and Question Design.” *HowTo.gov*, April 23, 2013. Accessed February 3, 2014. <http://www.howto.gov/customer-experience/collecting-feedback/basics-of-survey-and-question-design>.
- Flanagan, Mary, and Peter Carini. “How Games Can Help Us Access and Understand Archival Images.” *American Archivist* 75, no. 2 (October 1, 2012): 514–37. Accessed April 7, 2014. <http://archivists.metapress.com/content/B424537W27970GU4>.

- Frevert, Rhonda Huber. "Archives Volunteers: Worth The Effort?" *Archival Issues* 22, no. 2 (1997): 147–62. Accessed February 28, 2016. <http://www.jstor.org/stable/41101978>.
- Gilliland-Swetland, Luke J. "The Provenance of a Profession: The Permanence of the Public Archives and Historical Manuscripts Traditions in American Archival History." *The American Archivist* 54, no. 2 (1991): 160–75. Accessed February 19, 2016. <http://www.jstor.org/stable/40293549>.
- Goetz, Thomas. "Open Source Everywhere." *Wired* 11, no. 11 (November 2003). Accessed October 20, 2013. http://www.wired.com.ezproxy.mtsu.edu/wired/archive/11.11/opensource_pr.html
- Greene, Mark A., and Dennis Meissner. "More Product, Less Process: Revamping Traditional Archival Processing." *The American Archivist* 68, no. 2 (2005): 208–63. Accessed February 12, 2016. <http://www.jstor.org/stable/40294290>.
- Harsanyi, David. "The Amateurs' Hour: Is the Internet Destroying Our Culture, Or Is It Just Annoying Our Snobs?" *Reason*, 2008. Accessed February 6, 2016. <https://ezproxy.mtsu.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=edsgao&AN=edsgcl.172291492&site=eds-live&scope=site>.
- Holley, Rose. "Crowdsourcing: How and Why Should Libraries Do It?" *D-Lib Magazine* 16, no. 3/4 (April 2010). Accessed November 3, 2013. doi:10.1045/march2010-holley.
- Hustinx, Lesley, Ram A. Cnaan, and Femida Handy. "Navigating Theories of Volunteering: A Hybrid Map for a Complex Phenomenon." *Journal for the Theory of Social Behaviour*, 40, no. 4 (2010): 410–434. Accessed August 22, 2014. doi:10.1111/j.1468-5914.2010.00439.x.
- Johnson, Catherine A., and Wendy M. Duff. "Chatting up the Archivist: Social Capital and the Archival Researcher." *The American Archivist* 68, no. 1 (2005): 113–29. Accessed February 19, 2016. <http://www.jstor.org/stable/40294259>.
- Keen, Andrew. *The Cult of the Amateur: How Blogs, MySpace, YouTube, and the Rest of Today's User-Generated Media Are Destroying Our Economy, Our Culture, and Our Values*. New York: Random House, 2008.
- Keohane, Kevin. "Unpopular Opinion: Everyone's An Expert on the Internet. Is That Such a Bad Thing?" *Communication World*, 2008. Accessed February 6, 2016. <https://ezproxy.mtsu.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=edsgao&AN=edsgcl.173021681&site=eds-live&scope=site>.
- Krause, Magia Ghetu, and Elizabeth Yakel. "Interaction in Virtual Archives: The Polar Bear Expedition Digital Collections Next Generation Finding Aid." *The American Archivist* 70, no. 2 (2007): 282–314. Accessed February 11, 2016. <http://www.jstor.org/stable/40294572>.

- Lee Ann Potter, and Rebecca Martin. "NARA's Armies of Volunteers." *Prologue* 38, no. 4 (Winter 2006). Accessed January 13, 2016.
<https://www.archives.gov/publications/prologue/2006/winter/volunteers.html>.
- Leland, Waldo. "The First Conference of Archivists, December 1909: The Beginnings of a Profession." *American Archivist* 13, no. 2 (April 1, 1950): 109–20. Accessed April 7, 2014. doi:10.17723/aarc.13.2.h874j87h80441422.
- Leonard, Kevin B. "Volunteers in Archives: Free Labor, But Not Without Cost." *Journal of Library Administration* 52, no. 3/4 (April 2012): 313–20. Accessed February 19, 2016. doi:10.1080/01930826.2012.684529.
- McCune, Bonnie F. "The New Volunteerism: Making It Pay Off For Your Library." *American Libraries* 24, no. 9 (1993): 822–24. Accessed February 19, 2016.
<http://www.jstor.org/stable/25633036>.
- McHenry, Cheryl A. "Library Volunteers: Recruiting, Motivating, Keeping Them." *School Library Journal* 34, no. 9 (May 1988): 44. Accessed February 19, 2016.
<https://ezproxy.mtsu.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=aph&AN=5770264&site=eds-live&scope=site>.
- Nicol, Erica A., and Corey M. Johnson. "Volunteers in Libraries: Program Structure, Evaluation, and Theoretical Analysis." *Reference & User Services Quarterly* 48, no. 2 (Winter 2008): 154–63. Accessed April 7, 2014.
<https://ezproxy.mtsu.edu:3443/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=aph&AN=35665049&site=eds-live&scope=site>.
- O'Connell, Brian, ed. *America's Voluntary Spirit: A Book of Readings*. New York: Foundation Center, 1983.
- Owens, Trevor. "Digital Cultural Heritage and the Crowd." *Curator: The Museum Journal* 56, no. 1 (January 1, 2013): 121–30. Accessed May 24, 2013.
doi:10.1111/cura.12012.
- Parry, Marc. "Historians Ask the Public to Help Organize the Past." *Chronicle of Higher Education* 59, no. 2 (September 7, 2012): 17–17. Accessed September 8, 2015.
<http://chronicle.com/article/Historians-Ask-the-Public-to/134054/>.
- Pascoe, Peggy. *Relations of Rescue: The Search for Female Moral Authority in the American West, 1874-1939*. New York: Oxford University Press, 1990.
- Putnam, Robert D. *Bowling Alone: The Collapse and Revival of American Community*. New York: Simon & Schuster, 2000.
- Ransbotham, Sam, and Gerald C. Kane. "Membership Turnover and Collaboration Success in Online Communities: Explaining Rises and Falls from Grace in Wikipedia." *MIS Quarterly* 35, no. 3 (September 2011): 613–27. Accessed February 2, 2016.

<https://ezproxy.mtsu.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=bth&AN=63604897&site=eds-live&scope=site>.

- Saylor, Nicole, and Jen Wolfe. "Experimenting with Strategies for Crowdsourcing Manuscript Transcription." *Research Library Issues*, no. 277 (December 2011): 9–14. Accessed October 21, 2013. <http://publications.arl.org/rli277/10>.
- Schlitz, Stephanie A., and Garrick S. Bodine. "The Martha Berry Digital Archive Project: A Case Study in Experimental pEDagogy." *The Code4Lib Journal*, no. 17 (June 1, 2012). Accessed June 1, 2013. <http://journal.code4lib.org/articles/6823>.
- Shirky, Clay. *Cognitive Surplus: Creativity and Generosity in a Connected Age*. New York: Penguin Press, 2010.
- . *Here Comes Everybody: The Power of Organizing Without Organizations*. New York: Penguin Press, 2008.
- Stebbins, Robert A., and Margaret Graham, eds. *Volunteering As Leisure / Leisure As Volunteering: An International Assessment*. Cambridge: CABI Publishing, 2004.
- Stove, R.J. "The Cult of the Amateur: How Today's Internet Is Killing Our Culture." *National Observer*, 2007. Accessed February 6, 2016. <https://ezproxy.mtsu.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=edsgao&AN=edsgcl.173717050&site=eds-live&scope=site>.
- Tenopir, Carol. "Web 2.0: Our Cultural Downfall?" *Library Journal*, 2007. Accessed February 6, 2016. <https://ezproxy.mtsu.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=edsglr&AN=edsgcl.172905894&site=eds-live&scope=site>.
- Theimer, Kate. "What Is the Meaning of Archives 2.0?" *The American Archivist* 74, no. 1 (2011): 58–68. Accessed February 19, 2016. <http://www.jstor.org/stable/23079001>.
- Tocqueville, Alexis de. *Democracy in America*. Translated by Arthur Goldhammer. Library of America 147. New York: Penguin Putnam, 2004.
- Van Ness, Carl. "Much Ado about Paper Clips: 'More Product, Less Process' and the Modern Manuscript Repository." *The American Archivist* 73, no. 1 (2010): 129–45. Accessed February 12, 2016. <http://www.jstor.org/stable/27802718>.
- Wade, Louise Carroll. "Settlement Houses." *Electronic Encyclopedia of Chicago*, 2005. Accessed February 18, 2016. <http://www.encyclopedia.chicagohistory.org/pages/1135.html>.
- Wyman, Bruce, Rich Cherry, Doug Hiwiller, Jennifer Trant, and Susan Chun. "Steve.museum: An Ongoing Experiment in Social Tagging, Folksonomy, and Museums." In *Museums and the Web 2006: Proceedings*, edited by J. Trant and

D. Bearman. Toronto: Archives & Museum Informatics, 2006. Accessed January 22, 2016. <http://www.archimuse.com/mw2006/papers/wyman/wyman.html>.

Yeung, Anne Birgitta. "The Octagon Model of Volunteer Motivation: Results of a Phenomenological Analysis." *Voluntas: International Journal of Voluntary & Nonprofit Organizations* 15, no. 1 (March 2004): 21–46. Accessed February 29, 2016. <https://ezproxy.mtsu.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=aph&AN=12817951&site=eds-live&scope=site>.

APPENDICES

APPENDIX A

“CROWDSOURCING: WHO VOLUNTEERS, AND WHY?”

What motivates you to participate in online crowdsourcing initiatives to index, tag, and transcribe historical documents and data? I would love to find out, so I wrote a survey! My name is Kayla Utendorf, and I am currently working on my thesis as part of the Masters in Public History program at Middle Tennessee State University. This anonymous survey should take no more than three minutes to complete. Responses to this survey will be used as part of my thesis on the topic of volunteers and their participation in crowdsourcing on academic websites. Through my research, I hope to better understand what groups tend to volunteer on these websites and what motivates them to volunteer.

Completing and submitting this survey indicates that you consent to participate in this project. Participation is voluntary, and withdrawal will not result in any penalty to the participant. You may withdraw from the survey by simply not submitting the form. This survey involves no more than minimal risk that would be faced in everyday activities. If you have any questions or comments concerning this survey or difficulties completing it, you may reach me at ku2c@mtmail.mtsu.edu. For additional information about giving consent or your rights as a participant in this study, please contact the MTSU Office of Compliance at (615) 494-8918. Thank you for your help!

1. What is your age?

- a. Under 18
 - b. 19-29
 - c. 30-39
 - d. 40-49
 - e. 50-59
 - f. 60-69
 - g. 70+
2. Do you currently live in the United States?
- a. Yes
 - b. No
3. What is your gender?
- a. Male
 - b. Female
 - c. Other (Please List)
 - d. Choose not to disclose
4. What is your race/ethnicity?
- a. White/Caucasian
 - b. Black/African American
 - c. Hispanic or Latino
 - d. Native Hawaiian or other Pacific Islander
 - e. Multi-Racial
 - f. Other (Please List)
 - g. Choose not to disclose
5. Are you currently employed? Check all that apply.
- a. Full-time
 - b. Part-time
 - c. Student
 - d. Retired
 - e. Unemployed
 - f. Choose not to disclose
6. What is your annual household income?
- a. Under \$20,000
 - b. \$20,000-\$29,999
 - c. \$30,000-\$39,999

- d. \$40,000-\$49,999
 - e. \$50,000-\$59,999
 - f. \$60,000-\$69,999
 - g. \$70,000-\$79,999
 - h. \$80,000-\$89,999
 - i. \$90,000-\$99,999
 - j. \$100,000-\$149,999
 - k. \$150,000-\$200,000
 - l. More than \$200,000
 - m. Choose not to disclose
7. How many hours a month to you spend volunteering on this website by transcribing/indexing documents, uploading/tagging photographs, proofreading text, etc.?
- a. Less than 5
 - b. 5-9
 - c. 10-14
 - d. 15-19
 - e. 20-24
 - f. 25-29
 - g. 30+
8. How did you find out about the website you volunteer with?
9. Please explain briefly what motivates you to volunteer your time on this website.
10. Do you volunteer your time on any similar websites?
- a. Yes (please list)
 - b. No

APPENDIX B

ADDITIONAL SURVEY RESULTS

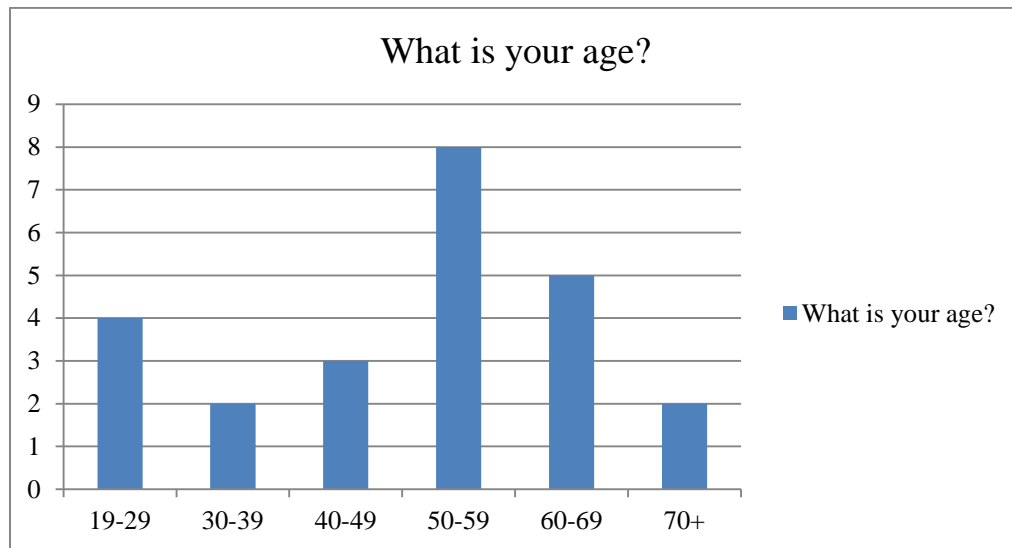


Figure 7. Age of survey respondents. Crowdsourcing: Who Volunteers, and Why? Survey by author.

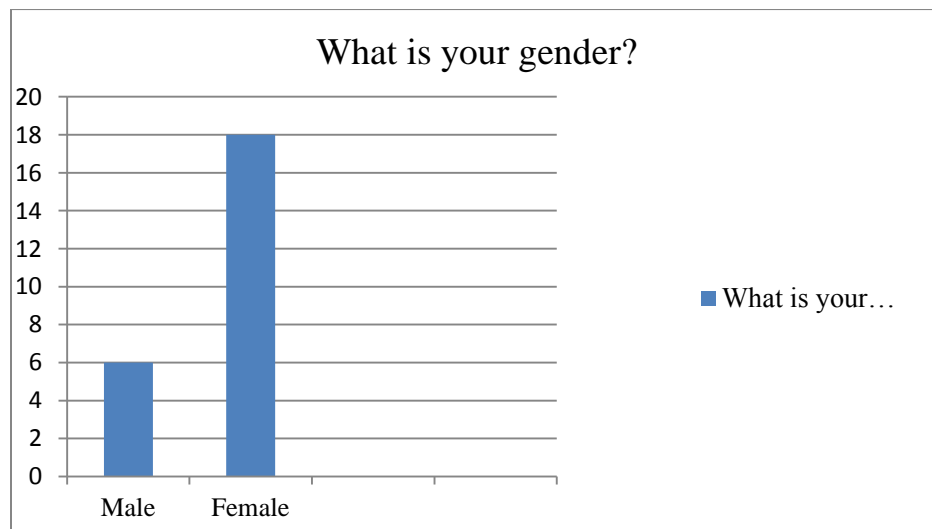


Figure 8. Gender of survey respondents. Crowdsourcing: Who Volunteers, and Why? Survey by author.

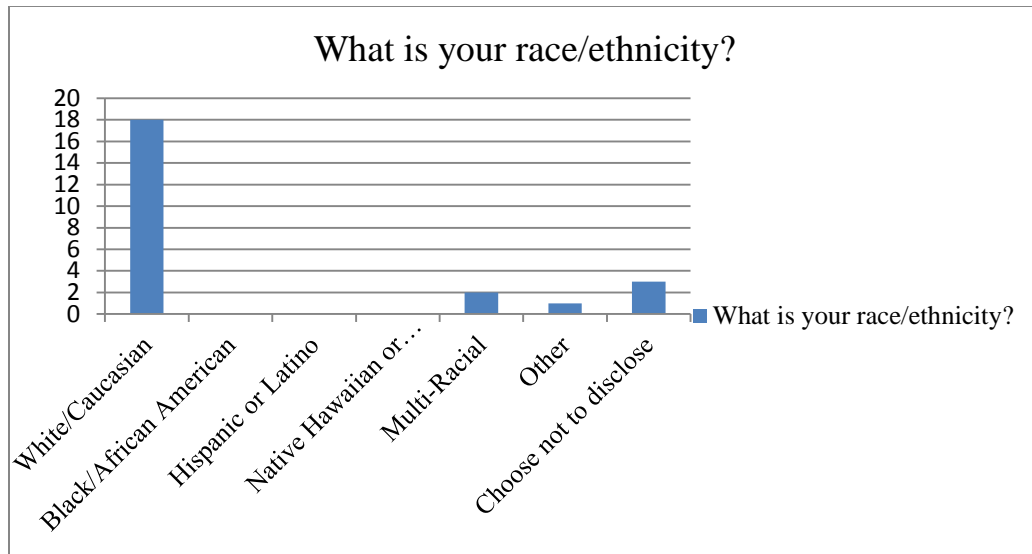


Figure 9. Ethnicity of survey respondents. Crowdsourcing: Who Volunteers, and Why? Survey by author.

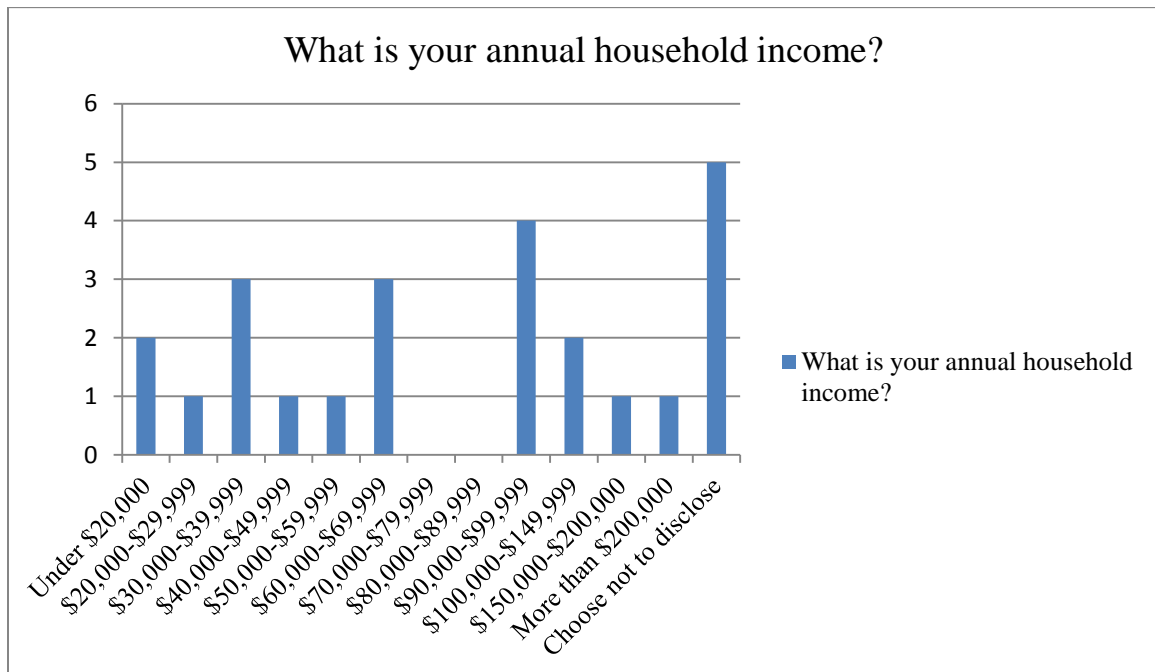


Figure 10. Annual household income of survey respondents. Crowdsourcing: Who Volunteers, and Why? Survey by author.

APPENDIX C

RECOMMENDATIONS FOR THOSE CONSIDERING CROWDSOURCING

What steps can be taken by those who are considering a new crowdsourcing project to help ensure the project's success? There are several recommendations those planning a crowdsourcing project should remember. Among those recommendations are to plan for all foreseeable outcomes, adapt your project for older adults who may not have strong technology skills, add some form of self-moderating scheme, provide adequate feedback, and provide forums for volunteers to communicate with each other.

Before beginning a crowdsourcing project, it is extremely important to *plan for all foreseeable outcomes*. Staff at the newly found Martha Berry Digital Archives comment that “healthy skepticism and astute design schemes are among our best defenses against the unruly flash mobs some critics of crowdsourcing fear.”¹ The authors’ reference to crowds as “flash mobs” makes a valid point—the public are not paid employees, and work solicited via crowdsourcing may be completed in bursts of activity, rather than following an orderly timeline and finishing by a previously mandated deadline. Institutions that need work completed by a specific deadline may find that crowdsourcing is not the right method for them. Staff at Institution A commented that anyone using volunteers should be “not too rigid in their standards and ideas and be willing to accept changes proposed by those volunteers.”² Staff interviewed at Institution C also recommended that anyone considering crowdsourcing examine all other similar

¹ Stephanie A. Schlitz and Garrick S. Bodine, “The Martha Berry Digital Archive Project: A Case Study In Experimental pEDagogy,” *The Code4Lib Journal*, no. 17 (June 2012).

² Manager at Institution A, email interview with the author, August 2014.

projects on the web. This way staff can agree on what ideas, methods, and interfaces they do and do not like, and by viewing the varying successes they can get an idea of what level of participation to expect with their own project.³

Planning for all foreseeable outcomes becomes especially important when planning a crowdsourcing project because one never knows how volunteers will react to a given project, or how staffing changes could affect the project. In 2005, a group at the University of Michigan began a project with the Bentley Historical Library and attempted to make an interactive finding aid for the Polar Bear Expedition Collections. Special tools on the website included the ability to bookmark pages, comment on items, “link paths” to similar records (for example “customers who viewed this item also viewed...”) and the option to make a user profile with biographical information.⁴ Researchers were disappointed, however, in the public’s use of the tools and concluded that the tools may not be the best for findings aids and archival interaction with patrons. In the six month test period, 114 visitors registered, but fifty-two of those registered never even logged on and only twelve actually participated in the website. Visitors only posted seventeen comments during the initial six month period, and most of those comments were intended for archivists, rather than for other visitors, as researchers had hoped.⁵ Eventually the students involved in the project moved on and the website became too difficult for the

³ Staff member at Institution C, Skype interview with the author, October 1, 2014.

⁴ Magia Ghetu Krause and Elizabeth Yakel, "Interaction in Virtual Archives: The Polar Bear Expedition Digital Collections Next Generation Finding Aid," *American Archivist*, 70 no. 2 (Winter 2007): 285-287, <http://www.jstor.org/stable/40294572>.

⁵ *Ibid.*, 312, 296-298, 310-311.

Bentley Historical Library to keep up to date, so the archives turned the site into a more traditional digital collection that is no longer interactive.⁶

In another example, The Bentham Project noted that it only had funding to provide full-time staff to the crowdsourced project for its six-month pilot period, and when they announced to volunteers that it would no longer be staffed with full-time employees, several of their regulars left the project out of disappointment. Bentham Project administrators felt this showed it would be impossible to completely get rid of staff moderation because volunteers needed encouragement and had questions that needed to be answered—any crowdsourcing project “must be based on mutual respect and trust” if it will escape accusations of being exploitative.⁷ The Bentham Project’s example also shows the foresight necessary on the part of those planning a project; if a project becomes too large for existing staff to manage and the institution has not prepared for the possibility of hiring extra staff, problems could be quick to follow. Making specific plans for various outcomes in advance can prevent headaches down the line. If the project receives less than stellar user engagement, will it be canceled, even if there are a handful of dedicated volunteers? Conversely, is the institution willing to hire more staff if the number of volunteers becomes unmanageable? Most archives can plan for steadier staffing than student interns. However, if only one staff member planned, implemented, and ran the project, who will manage the project if that person leaves? All these questions

⁶ “About the Polar Bear Expedition Digital Collections,” Polar Bear Expedition Digital Collections, accessed February 12, 2016, <http://quod.lib.umich.edu/p/polaread/about.html>.

⁷ Causer, et al, “Transcription Maximized, Expense Minimized?” 131-132.

and more should be considered before beginning a project to help prevent projects from fizzling down the road.

As noted earlier, *many crowdsourcing volunteers are older adults who may not have strong technology skills*. Archivists should be aware of this subset of volunteers to both recruit and retain members for their crowdsourcing projects by advertising the site where older adults are more likely to see it (perhaps mixing on- and off-line advertising) and by ensuring that crowdsourcing websites are easy to use for older adults who may not have as strong of technology skills as younger generations. Websites that require volunteers to encode their own submissions to achieve proper formatting may turn away otherwise eager participants, while websites set up like simple word processing software will likely be easier for the less tech-savvy.

Another recommendation to consider is to *add some type of self-moderating scheme*. Several archives have instituted such policies after getting their crowdsourcing efforts off the ground in attempts to lessen the amount of staff intervention necessary in editing volunteer transcriptions. There are several ways to introduce self-moderation. Institution A ranks their volunteers based on the quality of their work, and a single transcription is read multiple times by different volunteers, gradually moving up through the ranks for more and more detailed editing and will not be considered complete until the top tier of volunteer editors approves it.⁸ Old Weather is a crowdsourcing website which allows volunteers to transcribe climate data from old ship logs to be used by climatologists to document and predict climate change. Old Weather's system simply

⁸ Manager at Institution A, email interview with the author, August 2014.

requires several volunteers to transcribe a document and if all their notations agree, the document is assumed to be complete and correct.⁹

Institutions that allow users to self-moderate still need to *provide adequate feedback* to their volunteers so the volunteers are aware their work is appreciated. Few articles on traditional volunteering in archives discuss this (only “Using Volunteers for Special Project Staffing” from the National Air and Space Museum and the “Best Practices for Volunteers in Archives” from the Society of American Archivists even mention it)¹⁰, but nearly every institution felt that this was the most important recommendation to provide someone considering crowdsourcing. One author went so far as to remind readers to eventually add user-created tags to their institution’s metadata and to show the volunteers the tags had been added, so that volunteers knew their work meant something. Nicole Saylor, a staff member with DIY History, advised anyone beginning a project to recognize volunteers for their work (even going so far as to name individuals who had made large contributions) and to allow users to provide feedback in return. Interestingly, one DIY History volunteer commented on an online article about the website that he wished he had even more opportunities to communicate with staff and with other volunteers, so that volunteers could tutor each other and request specific feedback from staff. He commented that asking staff to research the name of a specific

⁹ “Frequently Asked Questions,” Old Weather, accessed September 11, 2015, <http://www.oldweather.org/faq>.

¹⁰ Susan E. Ewing, “Using Volunteers for Special-Project Staffing at the National Air and Space Museum Archives,” *The American Archivist* 54 no. 2 (1991): 183, accessed January 23, 2016, <http://www.jstor.org/stable/40293550>; Society of American Archivists, “Best Practices for Volunteers in Archives,” *Society of American Archivists*, August 2014, <http://www2.archivists.org/standards/best-practices-for-volunteers-in-archives>.

store he was finding regularly in a diary he was transcribing would likely require less effort than if they had to correct his mistake repeatedly throughout the course of the transcribed diary.¹¹ Institution A recommended to “think of it as a relationship, not a transaction, recognize [the volunteers] are human beings, and respect the opportunity and their knowledge.”¹² Lastly, Institution B mentioned to be aware of the *type* of feedback staff offers—noting that right and wrong answers should be clear and that inconclusive feedback can be frustrating.¹³ For example, when transcribing documents, the right answer is clear—volunteers can feel assured that if they are transcribing exactly what they see, their answers are correct. When providing metadata, however, volunteers require strong feedback to know whether to provide phrases or single words, synonyms, whether to capitalize words, or anything else on a long list of questions that could affect the quality of the metadata.

Lastly, archivists should *consider providing forums or some other communication scheme* to allow volunteers to develop a sense of community. By allowing volunteers to communicate with each other, crowdsourcing becomes more similar to traditional volunteering in that volunteers are able to develop strong friendships. Both traditional and surveyed crowdsourcing volunteers are strongly motivated to continue their work by

¹¹ Cairns, “Mutualizing Museum Knowledge,” 113; Owens and LeFurgy, “Crowdsourcing the Civil War,”; David Paul Davenport, comment on Jie Jenny Zou, “Civil War Project Shows Pros and Cons of Crowdsourcing,” *The Chronicle of Higher Education*, June 14, 2011, accessed September 10, 2015, <http://chronicle.com/blogs/wiredcampus/civil-war-project-shows-pros-and-cons-of-crowdsourcing>.

¹² Staff member at Institution C, Skype interview with the author, October 1, 2014.

¹³ Staff member at Institution B, telephone interview with the author, November 6, 2014.

the friendships they develop, and institutions should seek to foster these friendships by providing forums. As one surveyed volunteer commented “[Volunteering] is such a part of my life—it would be like walking away from family.”¹⁴ Volunteers and institutions both stand to benefit when volunteers are able to form a strong community.

¹⁴ Anonymous respondent, “Crowdsourcing: Who Volunteers, and Why?” Internet Survey, April 8, 2014.

APPENDIX D
IRB APPROVAL LETTER FOR SURVEY



3/18/2014

Investigator(s): Kayla Utendorf, Kelly Kolar
Department: Public History
Investigator(s) Email Address: ku2c@mtmail.mtsu.edu, kelly.kolar@mtsu.edu

Protocol Title: Crowdsourcing: Who Volunteers, and Why?

Protocol Number: #14-277

Dear Investigator(s),

Your study has been designated to be exempt. The exemption is pursuant to 45 CFR 46.101(b)(2) Educational Tests, Surveys, Interviews, or Observations.

We will contact you annually on the status of your project. If it is completed, we will close it out of our system. You do not need to complete a progress report and you will not need to complete a final report. It is important to note that your study is approved for the life of the project and does not have an expiration date.

The following changes must be reported to the Office of Compliance before they are initiated:

- Adding new subject population
- Adding a new investigator
- Adding new procedures (e.g., new survey; new questions to your survey)
- A change in funding source
- Any change that makes the study no longer eligible for exemption.

The following changes do not need to be reported to the Office of Compliance:

- Editorial or administrative revisions to the consent or other study documents
- Increasing or decreasing the number of subjects from your proposed population

If you encounter any serious unanticipated problems to participants, or if you have any questions as you conduct your research, please do not hesitate to contact us.

Sincerely,

Kellie Hilker, Compliance Officer
Office of Compliance
615-494-8918

APPENDIX E
IRB APPROVAL LETTER FOR INTERVIEWS



7/2/2014

Investigator(s): Kayla Utendorf, Dr. Kelly Kolar
Department: Public History
Investigator(s) Email Address: ku2c@mtmail.mtsu.edu; kelly.kolar@mtsu.edu

Protocol Title: Researching the Productivity of Crowdsourcing Projects

Protocol Number: #15-001

Dear Investigator(s),

Your study has been designated to be exempt. The exemption is pursuant to 45 CFR 46.101(b)(2) Educational Tests, Surveys, Interviews, or Observations.

We will contact you annually on the status of your project. If it is completed, we will close it out of our system. You do not need to complete a progress report and you will not need to complete a final report. It is important to note that your study is approved for the life of the project and does not have an expiration date.

The following changes must be reported to the Office of Compliance before they are initiated:

- Adding new subject population
- Adding a new investigator
- Adding new procedures (e.g., new survey; new questions to your survey)
- A change in funding source
- Any change that makes the study no longer eligible for exemption.

The following changes do not need to be reported to the Office of Compliance:

- Editorial or administrative revisions to the consent or other study documents
- Increasing or decreasing the number of subjects from your proposed population

If you encounter any serious unanticipated problems to participants, or if you have any questions as you conduct your research, please do not hesitate to contact us.

Sincerely,

Lauren K. Qualls, Graduate Assistant
Office of Compliance
615-494-8918

APPENDIX F
INFORMED CONSENT FOR INTERVIEWS

Informed Consent

Middle Tennessee State University

Project Title: Researching the Productivity of Crowdsourcing Projects

Purpose of Project: To discover the benefits and drawbacks, especially in terms of productivity, of crowdsourced projects. I hope to discover whether volunteers produce enough work of adequate quality to rationalize the extra work they create for those who must manage them.

Procedures: After returning this form, you will receive a list of questions via email. You may choose to answer those questions through email or through a telephone interview or Skype interview to be set up at your and the investigator's mutual convenience.

Risks/Benefits: This project should include no more risk than that which is encountered in everyday life. The results from this study will help us understand the usefulness of crowdsourcing volunteers.

Confidentiality: For the purpose of this study, your responses will be reported under the name of a pseudonym to retain anonymity.

Principal Investigator/ Contact Information: Kayla Utendorf, ku2c@mtmail.mtsu.edu

Participating in this project is voluntary, and refusal to participate or withdrawing from participation at any time during the project will involve no penalty or loss of benefits to which you might otherwise be entitled. All efforts, within reason, will be made to keep the personal information in your research record private but total privacy cannot be promised, for example, your information may be shared with the Middle Tennessee State University Institutional Review Board. In the event of questions or difficulties of any kind during or following participation, you may contact the Principal Investigator as indicated above. For additional information about giving consent or your rights as a participant in this study, please feel free to contact the MTSU Office of Compliance at (615) 494-8918.

Consent

I have read the above information and my questions have been answered satisfactorily by project staff. I believe I understand the purpose, benefits, and risks of the study and give my informed and free consent to be a participant.

SIGNATURE

DATE

APPENDIX G

COPYRIGHT PERMISSIONS

To Whom It May Concern:

I am completing a master's thesis on volunteering and crowdsourcing at Middle Tennessee State University. I would like to request permission to use a screenshot from the Galaxy Zoo website in my thesis. I have attached a copy of this screenshot for your perusal.

The thesis will be made available to the public through Middle Tennessee State University's library. In addition, the thesis will be available on the Internet through ProQuest. If this arrangement is acceptable, please supply me a signed letter granting me permission to use the screenshot. You can return the letter to me by email at this address or by mail at:

Kayla Utendorf
1510 Twin Oaks Dr.
Toledo, OH
43615

Thank you for your help.

Sincerely,

Kayla Utendorf

PERMISSION GRANTED:

I have the authority to grant the permission requested herein and I hereby grant permission to use the above referenced material in the manner described.

I request that the credit line read:

GALAXY ZOO IS A ZOOUNIVERSE.ORG PROJECT

Signature: 

Name & Title: DR GRANT MILLER

Institution: UNIVERSITY OF OXFORD / ZOOUNIVERSE

Date: 2/3/2016

To Whom It May Concern:

I am completing a master's thesis on volunteering and crowdfunding at Middle Tennessee State University. I would like to request permission to use a screenshot from the DIY History website in my thesis. I can send a copy of the screenshot for your approval.

The thesis will be made available to the public through Middle Tennessee State University's library. In addition, the thesis will be available on the Internet through ProQuest. If this arrangement is acceptable, please supply me a signed letter granting me permission to use the screenshot. You can return the letter to me by email at kn2c@utmail.mtsu.edu or by mail at:

Kayla Utendorf
1510 Twin Oaks Dr.
Toledo, OH
43615

Thank you for your help.

Sincerely,

Kayla Utendorf

PERMISSION GRANTED:

I have the authority to grant the permission requested herein and I hereby grant permission to use the above referenced material in the manner described.

I request that the credit line read:

University of Iowa Libraries

Signature: Matthew Butler

Name & Title: Matthew Butler, Senior Developer

Institution: University of Iowa Libraries

Date: 3/2/2016

To Whom It May Concern:

I am completing a master's thesis on volunteering and crowdsourcing at Middle Tennessee State University. I would like to request permission to use a screenshot from the Transcribe Bingham website in my thesis. I have attached a copy of this screenshot for your perusal.

The thesis will be made available to the public through Middle Tennessee State University's library. In addition, the thesis will be available on the Internet through ProQuest. If this arrangement is acceptable, please supply me a signed letter granting me permission to use the screenshot. You can return the letter to me by email at this address or by mail at:

Kayla Giendorf
1513 Twin Oaks Dr.
Toledo, OH
43615

Thank you for your help.

Sincerely,

Kayla Giendorf

PERMISSION GRANTED:

I have the authority to grant the permission requested herein and I hereby grant permission to use the above referenced material in the manner described.

I request that the credit line read:

COURTESY OF UCL SPECIAL COLLECTIONS

Signature: LS

Name & Title: DR. LOUISE SCARFORD, RESEARCH ASSOCIATE

Institution: BENTLEY PROJECT, UNIVERSITY COLLEGE LONDON

Date: 02/02/2016