3-25-2014

A History of Collaboration, a Future in Crowdsourcing: Positive Impacts of Cooperation on British Librarianship

Sally Ellis
Claremont University Consortium

Sally Ellis

A History of Collaboration, a Future in Crowdsourcing: Positive Impacts of Cooperation on British Librarianship

Abstract: This article looks at the varying definitions and usage of the term 'crowdsourcing,' including those that insist the term may only be applied to online activities. An argument is made that, semantics aside, the activities known as crowdsourcing, collaboration, and/or citizen science, were in use long before the current online medium was invented and should be recognized not for the terminology employed but for the mutual benefits reaped. The article addresses the strengths of such activities, including user engagement, the accomplishment of tasks otherwise not possible because of budget cuts, and diversity of viewpoint. In addition, the possible weaknesses of recruitment, perceived loss of power by professionals, and oversight of this method of project management are reviewed and suggestions for mitigation are proposed. Finally, instances of historic and contemporary ventures in British librarianship including the creation of the Oxford English Dictionary, the Mass Observation Movement, Oxford University's Ancient Lives project, FamilySearch genealogical/archives work, the British Museum and iGLAM collaboration with Wikipedia, and the Sounds of the U.K. are profiled. Arguments for the relative merits of each profile are highlighted, including benefits to society as a whole, the specific institution, its users, and the future of cultural heritage organizations throughout the United Kingdom.

Sally Ellis: M.S., Academic Librarian
Email: sallyellis19@gmail.com

Introduction

"This is my truth, tell me yours" – Aneurin Bevan[1]

British librarianship has been and continues to be strengthened by cooperation with, and contributions from, users; relationships among libraries, archives, museums (LAM) and other information institutions; and the use of emerging technologies to facilitate these associations and interactions. This paper will look at the various types of collaborations implemented both historically and presently. It will show that regardless of format and the varying semantic differences, collaboration between institutions and users – 'crowdsourcing' – solves problems, strengthens collections and communities, and engages users, thereby ensuring a future for libraries, museums, and archives in Britain.

Definitions of Crowdsourcing

Crowdsourcing is a newer term, so new in fact that some dictionaries– including the *Oxford English Dictionary* ("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.")— have only recently recognized its standardized use and created an entry for it. Jeff Howe, writing for *Wired* in 2006, coined the term 'crowdsourcing' and applied the expression to a number of different activities. Howe gave the example of the non-profit/public sector National Health Museum opting to use stock photos available from iStockphoto rather than commissioning photographs from a single photographer for educational kiosks on world health topics. iStockphoto grew out of a network of graphic designers who collected and shared their images with each other freely. In

---

[1] Aneurin "Nye" Bevan (1897–1960) was a British Labour Party politician, and a Member of Parliament from 1929 to 1959. A 'conviction politician,' he would often end his speeches with this phrase to express his belief that there is no universal truth, but only personal truths that we must respect.

2006, that network had grown to 22,000 contributors and included amateurs and professionals alike. Images had a price point of $1–$5 each, compared to $100–$150 that a single professional photographer might charge a non-profit for commissioned work. Another example described the 'outsourcing' of R&D ideas to a large network of engineers, scientists, and tinkerers by large corporate bodies such as Boeing, DuPont, and Procter & Gamble, whose in-house R&D (research and development) departments were floundering in an attempt to keep up with fast paced change while simultaneously cutting costs. Lastly, Amazon's Mechanical Turk was described as using its existing network to "find people to perform tasks computers are generally lousy at – identifying items in a photograph, skimming real estate documents to find identifying information, writing short product descriptions, transcribing podcasts" (Howe 2006, 4). The micropayments Amazon paid (and continues to pay) to its 'crowdsourcers' allows them to accomplish a great deal more than they would have if they were to pay a single full-time employee a living wage. In every example given, the use of large collective networks of both professionals and amateurs alike had effectively undercut the individual in price, innovation, and time savings.

Howe (2006) never strictly defined crowdsourcing, although implicit in his examples were some defining criteria: large networks of amateurs and professionals generating content, performing research, innovating ideas or performing tasks; that talent and/or those resources being sought either by companies or collectively pooled by participants in some sort of networked fashion to be made use of, either for free, for micropayments, or for larger sums (but still less than the hiring of a full-time employee). Brabham (2010) contends that only *online* projects should be termed 'crowdsourcing,' and a discussion and comparison of collaboration and 'crowdsourcing' would not be complete without acknowledging the lack of consensus among users of the term (hereafter used without explicit quote marks but with implicit marks, conceding that each reader may have a differing opinion as to the definition of the word). Without an authoritarian definition, however, the term has been applied to a number of different projects and has been used in diverse ways to describe the process of using large groups of people to meet a need, either through user-generated (UG) content, UG research, transcription, editing, or, at the IdeaStore in London, incorporating the life experiences and skills of users into the library's 'collection' via programming offerings in order to create a more user-friendly library experience.[2]

In March 2012, Estellés-Arolas and González-Ladrón-de-Guevara (2012, 197) published the results of their survey of the use of the term crowdsourcing in science-based journals in the Journal of Information Science. They looked at the most cited and prolific authors on the topic (published between 2006 and 2011) as well as subsequent articles that referenced back to the original texts, and undertook an examination of common and disparate elements of definitions in order to come to a collective, consensus-based definition of the term. Individual authors sometimes had multiple definitions throughout their works. Estellés-Arolas and González-Ladrón-de-Guevara's defining characteristics, which must all be met in order to qualify, are as follows:

– "(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 participative type;
– (g) it uses an open call of variable extent;
– (h) it uses the Internet." (2012, 197)

They then charted several projects which are commonly referred to as crowdsourcing and rated them against these criteria. Many projects did not fit the definition, most notably Wikipedia (Table 1).

It is important to note that Estellés-Arolas and González-Ladrón-de-Guevara (2012, 198), while surveying the literature available through Sage and Emerald, found a higher percentage of articles written for and by computer scientists. Further surveys of the literature more heavily weighted towards the "human sciences" and business may yield different results and may thus affect the continual evolution of the definition of crowdsourcing, especially given the number of varying applications termed 'crowdsourcing' within those disciplines.

What becomes obvious while reading this study is the indeterminate nature of the definition of crowdsourcing as used by both non-professionals and professionals. It could be argued that any one of the conclusions drawn is simply the opinion of the original author and serves his or her purposes and is then perpetuated by subsequent citation, whether "correct" or not. Given that, by its very nature, there is no single authority on the Internet and all its various activities, his conclusions here may not

---

2 The IdeaStore (http://www.ideastore.co.uk/en/home) is a public library in the Whitechapel neighbourhood of east London. It serves an urban, immigrant population using innovative service-oriented, patron-centric strategies.

Table 1: Conformity with the definition of crowdsourcing. Adapted from Estellés-Arolas and González-Ladrón-de-Guevara (2012, 197).

Criteria	a	b	c	d	e	f	g	h
Project								
Amazon Mechanical Turk	√	√	√	√√	√	√	√	√
Delicious	√							√
Fiat Mio	√	√	√	√	√	√	√	√
Flickr	√			√				√
InnoCentive	√	√	√	√	√	√	√	√
iStockphoto	√	√	√	√	√	√	√	√
Lánzanos	√	√	√	√	√	√	√	√
ModCloth	√	√	√	√	√	√	√	√
Threadless	√	√	√	√	√	√	√	√
Wikipedia	√	√	√			√		√
YouTube	√							√

apply to the world of Libraries, Archives, and Museums (LAM).

The distinction as it relates to the past and future of the library, archives, and museum world may be moot, however. What seems obvious from this study is that if strict definitions of crowdsourcing are going to be applied, at the very least, names for the other activities that don't follow that definition need to be given official recognition by determining and agreeing to suitable term(s). Or more likely, the recognition of the fact that crowdsourcing is only different from 'collaboration,' 'peer production/participation,' 'citizen science' by virtue of an online format. This becomes especially true after learning of the history of incorporating users' knowledge within information and cultural memory institutions – projects that have made the transition from offline to online medium and have stayed the same otherwise. Because clearly, without Estellés-Arolas and González-Ladrón-de-Guevara's (2012) strict definition, or if the defining characteristic of online medium were eliminated, it is obvious that the concept of "... outsourcing tasks, traditionally performed by an employee or contractor, to an undefined, large group of people or community (a 'crowd'), through an open call" has been around a very long time within many sectors of society: the arts, humanities, sciences and social sciences. It is only the *format* by which that call is announced and answered that has changed (Law and von Ahn 2011, 5).

For the purposes of this paper, then, crowdsourcing will be defined as a business model used by for-profit, not-for-profit, governmental or individual sectors to achieve a goal of collaboration, engagement, or task-oriented contribution by a large group of previous or new users in order to achieve results. The contributions may or may not be in exchange for payment and the medium for both the open call and the exchange of information is not confined to the Internet. It is true that the online format makes it easier to cast a wider net in the pool of users, as well as streamlines the transmission of information and makes it easier to manipulate in its digital format. The end result, however, remains the same.

Strengths and Weaknesses of User Collaboration and Crowdsourcing

The benefits of using collaboration both off and online are manifold in the library, archive, and museum and world. As Surowiekci (2005, xiii) states in *Wisdom of Crowds: Why the Many Are Smarter Than the Few and How Collective Wisdom Shapes Business, Economies, Societies and Nation,* "under the *right circumstances*, groups are remarkably intelligent, and are often smarter than the smartest people in them" [emphasis added]. There are certainly verifiable instances where groups have acted, collectively, less intelligently than an individual might have. However, crowdsourcing, with its diverse users, decentralization of power, ability of individuals to make independent decisions (and an aggregator who takes varying opinions and comes to a collective decision), is an ideal mechanism for harnessing the power of a crowd to make a good decision (Surowiekci 2005).

In Geneva, Switzerland (July 2010), a workshop was held to discuss the recent effective use of a widely dispersed crowd to perform evaluations on search results and the new methods for applying it.

> With regard to search evaluation, rather than employing in-house annotators for relevance assessment, one can instead leverage the "wisdom of the crowd" via this distributed workforce. While crowdsourcing poses a variety of new challenges in interacting with workers and ensuring standards

for quality control, a variety studies have shown that the crowd in aggregate can produce superior annotations to in-house assessors in less time and at significantly lower cost. (Carvahlo, Lease, and Yilmaz 2010, 17)

This collective intelligence has the benefit of raising the quality of the information provided. According to Holley (2010, 2–3), crowdsourcing accomplishes this in several ways: by the addition of tags which may be more diverse in nature than authority vocabulary dictates (which facilitate discovery by a diverse audience); through comments and ratings that give professionals insight into what users want to see and how they want to see it; by asking the crowd for answers to difficult questions (and then listening to the answers); and by adding value to the data through corrections, comments which contextualize the material (the modern day equivalent to marginalia), and tags. In the United States, one example of this type of value added is the Make History Project of the 9/11 Memorial Museum. Museum professionals have solicited stories, videos, and photographs of eyewitnesses to the destruction. The site's purpose is not only to gather narratives and artefacts of this time in history, but also to allow users to experience different perspectives of this day and to mini-curate those perspectives, by time or geography, using their own or other users' contributions. Each narrative or viewpoint, then, evolves into a new and possibly richer story when combined with others. In this manner, the site becomes not just a static collection, but also a 'shared narrative'; "it does not tell 'the' story of 9/11, instead it facilitates the sharing of the many stories of 9/11" (Murphy 2012, 12). As will be discussed later, some may view this ability by the crowd to add to or enhance museum or library materials as a threat and power shift from knowledgeable curators to an inconsistent crowd. Museum exhibition designer Nina Simon, however, contends "Many museums fear losing control... [but] [t]here's a difference between having power and having expertise... [m]useums will always have the expertise, but they may have to be willing to share the power" (Wright 2010, 1–2). Fergus Read (2012) of the Imperial War Museum understands this. The Museum uses offline partnerships with users, museum contributors, and their families to provide enhanced perspective and first person context to their artefacts through pictures, letters and audio recordings. To paraphrase Read (2012), having a gun or medal donated to the museum is a great thing; being able to attach a face and a story to that gun or medal is what makes for a powerful museum experience.[3]

Not only is the crowd able to accomplish these tasks *effectively*, but harnessing this aggregate intelligence is crucial to the *efficient* attempt to complete the backlog of projects that exist in cultural institutions. Huge amounts of material created too quickly to keep up with the organization of it, large volumes of collections that require tedious amounts of energy and focus, and/or the all too familiar refrain of budget and personnel cuts have caused these backlogs. One example of making use of the efficiency of crowdsourcing is the ReCaptcha project, which places words that were OCRed (optical character recognition) incorrectly or were unable to be OCRed, in 'captcha' windows on thousands of websites. When a sufficient number of users have transcribed the word, the system gathers that data and compares it in order to statistically determine the correct transcription (Parkes 2013, 66–69). In this way, users (most of them probably unaware that their typing is doing anything other than gaining them access to a website) are helping to digitize books and transcribe manuscripts in a way that will make more information available faster than any single librarian, archivist, or digitizer ever could.

Perhaps the benefit that is of greatest value to LAMs, however, is that of engaging existing and new users with library and cultural material. In an interview with advertising executive Evan Fry of Victor and Spoils, Fry describes the brand loyalty that crowdsourcing can bring to for-profit entities by "...using the power of the people and the magic that can bring to deliver the most relevant solutions where their customer base and culture at large feels ownership and love because they helped make it" (Liebling 2010, slide 14). It is a model that the cultural sector is wise to follow if they want to create a sense of pride and ownership in cultural and information institutions – ownership that instils a sense of responsibility for the continued financial viability and relevance of these libraries and museums. By seeking input from users and then listening and trusting that input to influence and shape collections, taxonomies, and user experiences, heritage and information institutions create a partnership, a trust, and a loyalty from old and new users alike. Building on past offline collaborative relationships while making use of today's online format provides for a more diverse and efficient relationship for growth and change. These symbiotic relationships will keep LAM organizations a vital part of the communities in which they reside (Holley 2010, 2–3).

This paper should in no way be construed as suggesting crowdsourcing as a panacea to all problems big and

---

3 For another project crowdsourcing war memorabilia and ephemera (this time digital objects only), see The Great War Archive

(http://www.oucs.ox.ac.uk/ww1lit/gwa), a collaboration between the Oxford University and JISC.

small in information and cultural institutions. There are definite weaknesses to the model that must be acknowledged, planned for, and mitigated, if a successful project is to be expected. Perhaps the biggest challenge exists on a systemic level. As alluded to previously, many librarians, archivists, and museum directors think of themselves as the gatekeepers of information – that the classification systems, ontologies, taxonomies, and provenance they assign to any item and to the information world at large is what keeps that world from descending into chaos. Inviting the public, both educated specialists and unvetted users, to create metadata, content, to transcribe historical documents or, in any way substitute their own expertise for that of the information professional, may be viewed as threatening to the experts' paradigm and certainly, at the very least, his livelihood. Holley (2010) counters that fear in her article "Crowdsourcing and Social Engagement: Potential, Power and Freedom for Libraries and Users." She argues that the world will always need libraries and librarians – it will need them to find, collect, describe, preserve, and offer up resources free of charge. We need not fear losing our power within the information arena. Given the consistent shortage of funds and workers, information professionals should harness new technologies and utilize crowdsourcing as a means to achieve the goals that professionals and users, alike, wish to achieve. They should allay their fears of loss of power in order to unleash greater power. Quoting Harriet Rubin, "Freedom is actually a bigger game than power. Power is about what you can control. Freedom is about what you can unleash" (Rubin 1998, as quoted by Holley 2010, 26–27).

On a much more practical level, crowdsourcing has some very real challenges. Every project will take oversight and money if it is to be successful. A project may not be offering any sort of payment for the work done, but money will still be necessary to support a paid employee as recruiter and manager of the project. Believing crowdsourcing to be simply a 'free' source of labour is erroneous. As with any face-to-face volunteer workforce, the crowd must be sought and the crowd and the project must be managed. The person in charge can affect both the input and output. Without strong leadership, vision, perseverance, and an encouraging nature, the project could go completely awry. Users must feel valued and engaged; they must see their work as making a difference and their time as well and interestingly spent, if they are to continue to contribute. Moreover, just as in face-to-face projects, virtual ones are subject to the same limitations based on personality, time management, and human error (Liebling 2010, slide 19).

Human error must definitely be taken into account when evaluating the time it will take to make these projects valuable and accessible to society. Typically, crowdsourced projects are those that a computer cannot or has not performed correctly. Therefore, by default, humans, in all their erring ways, may be the only alternative and answer for accomplishing a task. Rather than viewing human error as a weakness of crowdsourcing that takes time to correct, it should be remembered that, but for the humans and their efforts, the project or collection would remain 'dark,' of no use to anyone. Nevertheless, without doubt, one who specializes in the subject and is able to keep the existing historical standards of libraries, archives and museums, must indeed address quality control.

Another challenge lies in the use of user-generated metadata and tagging. Van Hooland, Mendez Rodriguez, and Boydens (2011, 709), discuss the 'commodification of culture,' a term which applies to the notion of placing extreme value on present accessibility of materials (as opposed to future accessibility ensured by authority control) because of the need to justify expenses and demonstrate 'fitness of purpose' as defined by ISO 2005 standards. In using a crowdsourcing method that incorporates user tagging and user-generated metadata to boost hits and access by current users, we must remember not to sacrifice our future users. Current users create social tagging within the context of the present. Left unchecked, however, those tags may be obsolete and irrelevant to users in the future. Employing user generated metadata and social tagging as a *complement* to authoritative vocabularies and established taxonomies can allay this concern. In tandem with information professionals, those UG tags and comments can open up collections and engage today's users by providing personal insights into items that were lacking in context and relevance; more structured metadata helps ensure that collections also continue to be accessible and valuable to future users.

One possible weakness that bears mentioning is the ethical dilemma of exploiting those who are so desperate for money or notoriety that they will give their labour away free or for pennies on the dollar. Eventually, this method of outsourcing tasks for micropayments could lead to increased supply of and decreased demand for paid, skilled labour. Additionally, some companies may pay micropayments for tasks and ideas, then turn around and make large profits on those initial miniscule investments. Within the private and for-profit sector, this may very well be a real concern. Aminda (2011) at IdeaConnection posits that, "Penny-pinching companies are hiring specialists to plumb the vast resources of the Web in search of cheap expert help." It is not within the scope of this paper to survey and determine to what extent this is a valid concern and business model within the private sector. It is, however,

possible to counter this concern within the non-profit realm with the argument that libraries and museums are uniquely positioned as 'public goods' and, as such, volunteers can be assured that their contributions will not be exploited for commercial gain (Holley 2010, 19).

Those who have never participated in either historical collaboration or online crowdsourcing may question why so many are willing to give freely of their time, their talents, and their labour in so many cultural and information endeavours. The answers are myriad. Wu (2005) explains one, in "the exposure culture...getting noticed is everything." Certainly, opportunities for notice exist in the crowdsourcing and social media movement.[4] Perhaps in the library and museum environment, however, this is not the primary incentive for users. Some of these reasons and the strengths and weaknesses of individual projects are analyzed in the following section.

Case Studies of Historical Collaborations and Current Crowdsourcing in the UK

Despite the drawbacks and with insight into the benefits, the following describe some of the projects which have used this collaborative/crowdsourced model. Some of the historical projects continue to exist and have successfully made the transition from paper to the online medium.

The Longitudinal Watch

In 1714, in an effort to keep sailors and shipping as safe as possible, and as a way to make empire seeking and building more probable, the British government offered a prize to anyone who could more easily establish a ship's longitude (latitude was established using the sun's altitude). The result was H4, a 'watch' designed by a working class joiner John Harrison. Over a 47-day voyage, the timepiece was accurate to within 39.2 seconds. It is available for viewing at the Royal Observatory today (O'Donnell 2002).

Oxford English Dictionary

In 1879, many years after its original announcement of plans to create an authoritative and complete chronological dictionary of the English language from Saxon times onward, the Philological Society of London appointed James Murray, a gifted linguist, to act as editor of the slow moving, sometimes stalled project. He was the third official editor, the first two serving short-term appointments due to illness or personality unsuitable to the scale and collaborative nature of the project. The original intent had been to publish a 6,400 page, four-volume dictionary based on the earliest uses of words found in literature. The project was ambitious and one man would not be able to complete it. Instead, it must be "the combined action of many." "It would be necessary to recruit a team – moreover, a huge one – probably comprising hundreds and hundreds of unpaid amateurs, all of them working as volunteers" (Winchester 1998, 106). Frederick Furnivall, a Society member who served as second editor, sent out a circular calling for volunteers to choose a period of history they would like to read from and send excerpts of word usage from that literature to his office on scraps of paper. The editor would then compile those papers alphabetically to form a chronological history of the defined uses of the word. In the end, the *Oxford English Dictionary* (OED) was born, using over six million reader-contributed scraps of paper defining 414,825 words in 1,827,306 literary illustrations (Winchester 1998, 109, 220). Today's users are invited to submit new words and usages for possible inclusion into the dictionary via online submission forms ("Rewriting the OED" 2012).

Mass Observation Movement

In 1937, three men who recruited volunteers to record their everyday observations in diary form or by answering open-ended questionnaires established the Mass Observation Movement in England. These public observers wrote down their thoughts about media, the actions of their neighbours, and generally created an "'anthropology of ourselves' – a study of the everyday lives of ordinary people in Britain." This study and its consequent reports, arguably, had some influence on life and political policy in Britain in the 1930's (Hubble 2010).[5] Students, research-

---

4 For a fun example of incentives, see the Old Weather Project (http://www.oldweather.org/), in which contributors work their way up the "ranks" from ship's Cadet to Lieutenant, to Captain, based on quantity of transcriptions completed. This project aims to use ships' logs to garner a more complete picture of weather patterns and historical ship voyages.

---

5 Another crowdsourced project that aims to influence public policy is the East London Legacy 2012 project (http://www.uel.ac.uk/ell2012/). This crowdsourced archive documents the impact of the 2012 London Olympics on the lives of East Londoners. Specifically, it

Table 2: FamilySearch Indexing - number of volunteers and achievements (Source: Holley 2010, 9)

Date	Number of volunteers	Cumulative achievements: Number of individual names transcribed
August 2005	FamilySearch Indexing on Web introduced.	
January 2006	2,004 online volunteers	
January 2007	23,000 online volunteers	102 million
January 2008		217 million
November 2009	160,000 online volunteers	334 million

ers, and teachers can now use this compilation as a means of understanding life in twentieth and twenty-first century Britain. The original project lasted until the late 1950s and was then revived in 1981. It now uses both postal and electronic mail (email) to convey its surveys to participants. The material is archived at the University of Sussex (Mass Observation 2012).

FamilySearch at the National Archives

While not strictly British, FamilySearch has made a home at the British National Archives and is contributing to British librarianship in a very real way. FamilySearch is a non-profit entity owned by the Church of Jesus Christ of Latter-Day Saint (LDS). Its mission and goal is to gather, photograph, transcribe, digitize, and make available as many internationally archived vital, civil, private, and religious records as possible. To accomplish this, Family-Search has solicited the help of many volunteers (see Table 2) (for example, 90,000 additional volunteers were recruited to transcribe the 1940 U.S. census). FamilySearch is a perfect example of historical collaboration that has transitioned into online crowdsourcing. Prior to the Web, FamilySearch used analogue formats (floppy disks, CDs, mail, etc.) to transmit the information between headquarters and volunteers. The introduction of the online medium has considerably speeded up the process of transcription and allows FamilySearch to open up the call to more volunteers worldwide.

In addition to transcription, FamilySearch also solicits user generated content in the form of family trees (family group sheets), published family and local histories, and archival records which may provide rich genealogical data.

As discussed earlier, crowdsourcing, while a boon to productivity, is not without challenges. One such challenge for human transcription, as is done with census re-

cords, is simple human error or differing opinions as to the reading of handwritten records. FamilySearch attempts to lessen these errors by requiring a double blind transcription process, much like the ReCaptcha program discussed earlier. Records are sent out to two different transcribers and are cross-checked for consistency. If both transcriptions do not match, a third person will attempt to reconcile the differences. If a suitable match is inconclusive, information specialists will seek more opinions. The same record may be sent out for additional opinions, if need be.

For user-generated content (family trees and linked connections), there are varying levels of commitment to accuracy, depending on the user. In an effort to differentiate "authoritative" records (those crowdsourced materials overseen by a trained specialist and ultimately uploaded by a FamilySearch employee) from user generated content (not verified by specialists and directly uploaded by users but of considerable potential value nonetheless), the latest update to the website more clearly separates the two. Indexed and transcribed records are available under 'Records' and UG material is grouped under 'Genealogies' when using the Search feature.

Genealogists are an interesting and engaged lot. Their commitment to finding information on their own relatives often translates into helping others find theirs. In this way, volunteers in crowdsourcing projects like FamilySearch are there because they are personally interested in the subject matter. (This is also true of those helping in the Transcribe Bentham Project.[6]) Additionally, because FamilySearch is tied to a religious mission of the LDS church, other volunteers see it as a service project that has immediate results and accomplishes their personal goal of helping the mission of the church prosper. Still others do it because they claim to learn new things each time.[7] This

---

looks at whether the Olympic bid promises of "Social regeneration" and "Health and wellbeing," were met.

6 University College London's 'Transcribe Bentham Project' (http://blogs.ucl.ac.uk/transcribe-bentham/) is a 'participatory project,' an online transcription of the papers of Jeremy Bentham, famed philosopher and reformer.

7 According to Stewart Gillies (2012), Information Services Manager at the British Library's national Newspaper Library, another project that benefits from knowledge thirsty volunteers

would be particularly true in the case of historical diaries (Holley 2010, 10–11).

By all accounts, FamilySearch in Britain is able to accomplish marvellous things through collaboration with users. All of this content benefits British librarianship, at large, which is quickly experiencing increasing demands on its archival content due to the burgeoning interest in family history. At this point, it is hard to tell which is driving which.

The British Museum and iGLAM Collaboration with *Wikipedia*

In 2010, British Museum (BM) website manager Matthew Cock (2010, 3) recognized a problem. For one of the Museum's most famous objects, the Rosetta Stone, BM website hits were 18,359 for the month of April. *Wikipedia*'s article on the Rosetta Stone was viewed five times more frequently (in English alone). Cock realized the potential for collaboration was there. While many would view the relationship between the two institutions as making for strange bedfellows (professional vs. amateur), Cock saw it differently: "...we need each other....we share some common goals." Among those goals is the desire to "build an [educational] resource that is made available to all the people of the world for free." The collaboration took the form of a "Wikipedian-in-Residence"— Liam Wyatt, a volunteer, spent 5 weeks building the relationship between the two institutions in order to work towards the goal of combining one's expertise and the other's established online audience. The result is articles about the items in the British Museum written by 40 volunteer *Wikipedia* editors and contributors with more intimate knowledge of the items and with a contextual understanding that the Museum's curators are most skilled and happy to give.

The crowdsourced collaboration was not without its trials. Point of view was one of them. While *Wikipedia* strives for a neutral point of view, often including multiple thoughts in an effort to remain neutral, the British Museum bases its point of view on individual academic scholarship, hoping that continued research will correct misconceptions and give greater perspective and new understanding to the knowledge base. The difference then is what could be termed the 'consensus view' versus the 'individual scholarship' viewpoint (Cock 2010, 12). The two processes of coming to distinct viewpoints can be meshed

if the goal of accurate information remains the same. In this manner, established consensus does not stifle the ability to learn and grow through ongoing scholarship. While Cock's main goal was not to drive more traffic to the BM website, some increased traffic has certainly occurred. More importantly, however, the collaboration has engaged volunteer *Wikipedia* editors in the Museum's content, made new editors out of BM curators, inspired people to talk (and write) about the British Museum, and widened the audience for all the amazing treasures the Museum holds. As Roger Pearse blogged "What is depicted is a model for institutions on how to deal with the Internet revolution. It's clever, it costs them nothing, it gains the institution respect and traction on the Internet... there is, in truth, no downside" (Pearse 2010, as quoted by Cock 2010, 26).

The Sounds of the UK

In mid-2010, the British Library, in cooperation with the Noise Futures Network, launched the UK SoundMap, a project whereby crowdsourced users could contribute sound bites of music, voices (including dialects and accents), narrations, wildlife recordings, atmospheric noise, and more to the Library's Sound and Vision Department database. The project is part of a larger project titled Unlocking and Integrating Audio Visual Content. The sounds are collected by crowdsourcees using the free Audiboo app, are geotagged, content tagged, and may be commented on by users. Library staff then adds an additional tag for use in RSS feeds and places the file on the SoundMap (using Google Maps) for a 'mash-up' curation of sounds across the UK. The project uses free, user-friendly technology (apps, uploading, etc.) available to Smartphone users and aims to accomplish three things:
–   "To explore the potential for mash-ups using data accumulated by cutting edge applications to build significant resources for digital scholarship
–   To map the evolution of the **national** landscape and record how people feel about it.
–   To involve the public in contributing to British Library acquisitions of research material." (Pennock and Clark 2011, 2)

During its twelve-month gathering period, 350 distinct users made over 2,000 contributions ("Sounds"). While officials deem the program a success according to its goals, it was not without some glitches. About 7% of the contributions were not added to the map. The reasons for this varied from untagged location, copyright infringement, poor

---

is the OCR transcription corrections made by those working with development of the British Newspaper Archive (http://www.britishnewspaperarchive.co.uk/).

audio quality, personal privacy issues, and obscenities, which were blocked because there was no way to warn listeners of its presence. Other potential risks involved reliance on outside technology (Audioboo and GoogleMaps) and possible abuses to the project by those seeking to exploit for "political gain" (Pennock and Clark 2011, 3). All of these risks were lessened by foreknowledge and expert handling by library specialists.

The SoundMap project is an example of an anonymous crowdsourced project; no user profiles were collected with the uploading of files. Therefore, ideas about crowdsourcee incentives or reasons for contribution are simply conjecture. This project, then, contradicts the notion of 'exposure culture' being the only incentive for contribution (Wu 2005). Some may cooperate for personal profit, some out of a desire to belong; others may simply do it for the benefit of humanity. Whatever the reason, the Soundmap created by the British Library and its crowdsourced users is a unique collection of the sounds and places of England in 2010 and 2011. It is available to scholar and non-scholar alike and is a great example of the use of emerging technologies for collaboration and crowdsourcing in libraries today.

Ancient Lives

Another British project taking advantage of the efficiency of large masses of people transcribing handwritten documents is the previously slow moving Oxford University's Ancient Lives project, which is using the term 'citizen science' to describe its collaborative nature. The collective of academics and "citizen scientists" is working to identify and transcribe papyri belonging to the Egypt Exploration Society and several other institutions. The process of uploading images and allowing crowdsourced transcribers and computer intelligence models to study papyri and make educated guesses as to the characters displayed there is "allowing professionals to process large batches of data at any given time... to identify known texts and documents faster than ever before (University of Oxford 2012).

Conclusion

Seeking the input, individual knowledge, and creative energy of library, museum, and archive users serves many purposes in the information world. Enriching collections, reducing backlogs of material, making the most of dwin-

dling budgets, finding solutions to problems, and engaging users, are just some of the benefits derived from collaborating and cooperating with others both in and outside the LAM discipline. This collaboration, or crowdsourcing, is not a new concept; it has been used in many offline projects for hundreds of years. Examples of this include the making of the *Oxford English Dictionary* and the Mass Observation Movement. Both of these projects were started in analogue format and both have made the transition to the online medium today. This medium makes collaboration easier and faster and makes the rich resources of libraries, museums, and archives more readily accessible to diverse users, myriad voices, and larger volunteer forces.

Though not without challenges as systemic as professional philosophies which include the concept of librarian as gatekeeper, and practical issues of volunteer recruitment and management, LAM's are unique in their ability to derive all that is good about crowdsourcing and mitigate much that is bad. Through continued use of this creative cooperation, the information world can build and strengthen relationships with engaged users who will, in turn, advocate for robust cultural and historical institution budgets. The relationship is a symbiotic one and, as more and more institutions make use of the growing technologies available to them, British librarianship will continue to experience growth and mine richer and richer content for today's and tomorrow's users.

References

Aminda. 2011. "Crowdsourcing: An Ethics Dilemma?" *IdeaConnection Blog*, 23 April. Accessed January 20, 2014. http://www.ideaconnection.com/blog/2011/04/crowdsourcing-an-ethics-dilemma.

Brabham, D. C. 2010. "Crowdsourcing as a Model for Problem Solving: Leveraging the Collective Intelligence of Online Communities for Public Good." PhD diss., University of Utah.

Carvalho, V. R., M. Lease, and E. Yilmaz. 2010. "Crowdsourcing for Search Evaluation." *SIGIR Forum* 44 (2): 17–22. doi: 10.1145/1924475.1924481.

Cock, M. 2010. "A Wikipedian-in-Residence at The British Museum." *Slideshare* 38 PowerPoint slides. Accessed July 8, 2012. http://www.slideshare.net/matthewcock/.

Estellés-Arolas, E., and F. González-Ladrón-de-Guevara. 2012. "Towards an Integrated Crowdsourcing Definition." *Journal of Information Science* 38 (2): 189–200. doi: 10.1177/0165551512437638.

Gillies, S. 2012. "Tour of British Newspaper Archive." Talk given at British Newspaper Archive, Colindale, UK, May 2012.

Holley, R. 2010. "Crowdsourcing: How and Why Should Libraries Do It?" *D-Lib Magazine* 16 (3/4): 5. Accessed June 15, 2012. http://www.dlib.org/dlib/march10/holley/03holley.html.

Howe, J. 2006. "The Rise of Crowdsourcing." *Wired* 14 (6). Accessed June 4, 2012. http://www.wired.com/wired/archive/14.06/crowds.html.

Hubble, N. 2010. *Mass Observation and Everyday Life: Culture, History, Theory*. New York: Palgrave Macmillan.

Law, E. L. M., and L. Von Ahn. 2011. *Human Computation*. San Rafael, CA: Morgan and Claypool.

Liebling, R. 2010. "Everyone is Illuminated: A Primer on Crowdsourcing – the Creative, Business and Social Implications of a Controversial Technique." *Slideshare* 54 PowerPoint slides. Accessed June 7, 2012. http://www.slideshare.net/eyecube/everyone-is-illuminated-3129260.

Mass Observation: Recording Everyday Life in Britain. 2012. "About Mass Observation." Accessed June 7, 2012. http://www.massobs.org.uk/menu_about_mass_observation.htm.

Murphy, O. 2012. "Crowdsourcing History." *SlideShare* 24 PowerPoint slides. Accessed June 10, 2012. http://www.slideshare.net/oonaghmurphy/crowdsourcing-lecture-pres.

O'Donnell, J. 2002. "John Harrison and the Longitude Problem: Royal Observatory and History of Astronomy." Royal Museums Greenwich. Accessed June 2012. https://web.archive.org/web/20120722015138/http://www.rmg.co.uk/harrison.

Parkes, D. C. 2013. "CS 186 Economics and Computation, Lecture 15: Human Computation and Crowdsourcing." Harvard School of Engineering and Applied Sciences. Accessed January 14, 2014. http://www.seas.harvard.edu/courses/cs186.

Pearse, R. 2010. "Wikipedia and the British Museum." *Roger Pearse: Thoughts on Antiquity, Patristics, Putting Things Online, Freedom Of Speech, Information Access, and More* (blog), June 8. Accessed on June 17, 2012. http://www.roger-pearse.com/weblog/2010/06/08/wikipedia-and-the-british-museum/.

Pennock, M., and C. Clark. 2011."Saving the Sounds of the UK in the UK SoundMap." *Ariadne* 66. Accessed June 9, 2012. http://www.ariadne.ac.uk/issue66/pennock-clark/.

Read, F. "Introductory Talk and Mini-Tour of Explore History Centre, Imperial War Museum." Talk given at Imperial War Museum, Department of Collection Access, London, UK, May 2012.

"Rewriting the OED." 2012. *Oxford English Dictionary*. Oxford University Press. Accessed July 26, 2012. http://www.oed.com/public/oed3/rewriting-the-oed.

Rubin, H. 1998. Answer to Question "How to Succeed at Business" at Business Roundtable Presentation for Fast Company, Telluride, CO, January 1998. Accessed July 3, 2012. http://www.fastcompany.com/33509/fast-pack.

Sounds. 2012. British Library. British Library Board. Accessed July 27, 2012. http://sounds.bl.uk/Sound-Maps/UK-Soundmap/.

Surowieckci, J. 2005. *The Wisdom of Crowds*. New York: Anchor Books.

University of Oxford. 2012. "Ancient Lives Research." *Zooniverse*. Accessed July 10, 2012. http://ancientlives.org/research.

Van Hooland, S., E. Mendez Rodriguez, and I. Boydens. 2011. "Between Commodification and Engagement: On the Double-Edged Impact of User-Generated Metadata within the Cultural Heritage Sector." *Library Trends* 59 (4): 707–720.

Winchester, S. 1998. *The Professor and the Madman: A Tale of Murder, Insanity, and the Making of the Oxford English Dictionary*. HarperPerennial ed. New York: HarperPerennial.

Wright, Alex. 2010. "Online, It's the Mouse That Runs the Museum." *New York Times.* January 19. Accessed July 6, 2012. http://www.nytimes.com/2010/01/20/arts/design/20museum.html?pagewanted=all&_r=0.

Wu, T. 2005. "Leggo My Ego: Google Print and the Other Culture War." *Slate*. Monday, Oct. 17. Accessed September 14, 2013. http://www.slate.com/articles/news_and_politics/jurisprudence/2005/10/leggo_my_ego.html.