

Navarrete, T. (2020) “Crowdsourcing the digital transformation of heritage” in Massi, Vecco and Lin (eds.) (2020) *Digital Transformation in the Cultural and Creative Industries*, London: Routledge, pp.99-116.

7 Crowdsourcing the digital transformation of heritage

By Trilce Navarrete

Abstract

Crowdsourcing, or outsourcing to the crowd, serves to redistribute the cost of developing new products and services beyond the firm to a crowd who provides finances, information, labor, or ideas for a marginal cost. Digital networked technologies have increased the pool of providers and speed of provision with important economic implications. Adoption in the heritage field has not been widespread. This chapter proposes a taxonomy of crowdsourcing projects and presents a selection of representative projects in the heritage sector focusing on advancing digital heritage activities. The analysis identifies four main benefits to involve the crowd: (1) work can advance surpassing limitations of institutional budgets constraints; (2) participation and further crowd engagement stimulate the notion of socially constructed commons; (3) tapping into a global crowd enables the redistribution of costs and benefits beyond geo-political boundaries, strengthening notions of a common heritage resource; and (4) by engaging in digital collaboration with the crowd, greater insight can be gained into the cultural consumption dynamics evolving on the Internet. Challenges identified include the sustainability of the crowd’s work, involving the reuse of results beyond the crowdsourcing project, as well as incorporating the crowd’s contribution into the organizational production process, adopting it for instance as part of the collections’ information system.

Keywords:

Crowdsourcing; Digital Heritage; Common Heritage; Sustainability

Introduction

Crowdsourcing can be defined as the redistribution of the cost of developing new products and services beyond the firm to a crowd who provides finances, information, labor, or ideas resulting in marginal costs for the firm (Howe, 2008; Estelles-Arolas & Gonzales-Ladrón-De-Guevara,

2012). This chapter looks at crowdsourcing and heritage that has been converted into a digital media format for transmission through the Web with the help of the crowd.

The World Wide Web was envisioned as a global document reference system, where individuals could navigate via hyperlinks to access various text, images, audio, and video in increasingly complex forms (Berners-Lee et al., 1992). Website design has developed to facilitate consumer creation and distribution of multimedia content (Visser, 2014), while access to lower-cost mobile technology has facilitated the increase in online use to 72.7% of the global population (ITU, 2018). It could be argued that digital transformation enables democratization of consumption and production of content to anybody with a device that can access the Web. Enclosed consumption (e.g., via a subscription of email, telephony, television, virtual private networks, gaming, and many mobile applications accessed without a browser) is outside the scope of this chapter, which focuses on accessing the Internet content through the web.

The original Web could have developed as a Peer-to-Peer network (P2P) but that required all users to be active editors and contributors. Instead, most consumers preferred the client-server model, which allows them to free-ride from the content produced by others (Kushner, 2016). Contributing to existing platforms represents a lower overall consumer's costs and so the Web 2.0 flourished, promising full democratization to consumers.

An explosion of projects followed which explored the production and distribution of content online involving the crowd, with strong network effects centralizing activity in a winner-takes-all model. Firms benefited from economies of scale, increasing the number of contributors, and from economies of scope, increasing the pool of ideas and general resources to develop new products. New digital platforms emerged to facilitate the interaction of actors, with perhaps the best-known being Amazon's Mechanical Turk, a platform that serves as marketplace for

initiators (crowdsources) to find workers to complete a task. Crowdsourcing became the term to refer to the outsourcing work to the virtual crowd (Estelles-Arolas & Gonzales-Ladrón-De-Guevara, 2012).

It can be argued that virtual work increases efficiency in the labor market by pairing actors, so that workers can participate from any location at any time, and that crowdsourcing may reduce brain drain (Cherry, 2011). However, few workers are paid, even though many receive some form of reward. The new technology has led to new dynamics in the use of resources calling for a revision of legislation: for instance, regarding virtual labor contracts and intellectual property rights. Something else that has changed is the ability to capture resources that used to be considered a public good, such as the knowledge about the world, and transforming them into commons triggering two social dilemmas: free-riding and sustainability (Bertacchini et al., 2012). This is the case of digital galleries, libraries, archives, and museums: the so-called GLAMs.

Digital technology facilitates the dissemination of heritage collections and public engagement. A sectoral goal, and no easy challenge, is to link digitized collections in meaningful structures across institutions and as part of the Web 3.0 network to position GLAMs prominently in the information economy (EC, 2011).¹ It has been estimated that 300 million objects from European heritage institutions have been digitized, representing 10% of the region's cultural heritage (Europeana, 2017). Identifying the provenance and context of heritage objects, for example, has not yet been automated and remains a high labor cost post. Crowdsourcing has attracted the attention of heritage institutions to advance this monumental task (Estermann, 2014). A few examples will be discussed in a later section. Engaging the public to advance the work of heritage institutions is not new to the adoption of digital technology. As example we can identify the Friends of the Museum association serving as financing crowd and donors providing the

majority of the museums' collections. Digital technology has, indeed, increased the speed of information spread and the scope of potential participants. Using social media, the Prado Museum has asked the online crowd to identify the location depicted in a painting recently donated.² At the time of writing, the museum is analyzing the numerous suggestions and images for possible identification.

In this contribution, a series of digital heritage projects that make use of crowdsourcing are analyzed to argue that the most successful projects are characterized by strong (self-)governance and niche markets. Examples presented in this chapter suggest that greater engagement with the content can be found in all activities of heritage institutions, where the crowd supports the management, preservation, and development of a cultural good. For specific examples, socially constructed commons can be identified.

The remainder of the chapter is organized as follows: Section Two will review the literature on crowdsourcing, presenting characteristics that will serve to form a typology to discuss the emerging dynamics. Section Three will focus on the heritage sector, particularly the process of digitization of content, and examples will serve to populate the typology constructed. Section Four will close with policy implications.

Crowdsourcing background

Crowdsourcing has been adopted as distributed solution to various problems with attractive cost savings for firms. A very early form of crowdsourcing can be found in the 16th century's rewards issued to devise a system to determine the location of a ship in the Netherlands (Bell, 1947). Online, the initiator, or crowdsourcer, sets out a call to the crowd to take part in a project, with various forms of contracts, providing certain rules and expectations for instance to specify if

the task is given to an individual or to the crowd (Berg et al., 2018) or regarding the ownership of the resulting product and intellectual capital generated (Wexler, 2010).

The crowdsourcing system is meant to distribute the risk involved when developing new products or services, and, therefore, the crowd may receive some form of payment. Payment for contributors of crowdsourcing varies and can be material or psychological (Wexler, 2010; Cherry, 2011). Individual ideas may involve a larger payment, while micro-tasks may be rewarded at a cent of a dollar per task (Hossain & Kauranen, 2015). Contributors may not always be paid in cash but may receive credit to purchase goods (Amazon's Mechanical Turk workers get paid in Amazon.com credit), virtual currency (common for videogames), or simply be given points to rank the level of achievement. In some cases, contributors participate in projects to gain experience, tap into a network, or receive career-enhancing visibility (Wexler, 2010).

Virtual labor has raised concerns regarding rights, safety, and payment of participants: for instance, to prevent discrimination based on age, gender, background, or location (Wexler, 2010; Cherry, 2011; Berg et al., 2018). Crowdsourced work is labor on demand, "temporary and very short-term," with extremely low wages if any (Ettliger, 2016, p. 2170). Certain crowdsourcing platforms specifically market their outsourced labor to the global south, advertising the possibility to advance social responsibility goals for potential firm clients (Prpić et al., 2015). An example is Samasource, a platform that offers training to the crowd arguing that outsourced virtual labor is a way to increase digital literacy in the global south (Prpić et al., 2015). Regions with low labor costs may be more prone to virtual work for micro-tasks, for example, for the gaming industry to advance the level of a player in a videogame. As crowdsourcing projects become gamified, the division of work and leisure may blur rising issues of responsibilities, ownership, and benefits.

Other forms of crowdsourcing may involve local governments seeking to stimulate public participation and dialogue through crowdsourcing projects, also known as citizen sourcing. Governments can collect ideas to solve problems, to develop policies, to design service co-production, or for surveillance and monitoring, and can also collect images and other documenting information (Nam, 2011). Examples include tracking the spread of influenza through a mobile phone application (Ranard et al., 2013) and the United Nations MY World survey and e-discussion on the topic of environmental sustainability (Gellers, 2016).³

A form of crowdsourcing often analyzed by economists is crowdfunding, which involves the collection of funds from the crowd. The main benefit of crowdfunding is overcoming the geographic limitations to identify potential investors (Robertson & Wooster, 2015), while the greatest challenge is that of asymmetric information that prevents backers from supporting a given project (Belleflamme et al., 2019). In the United States, the passing of the Jumpstart Our Business Startups (JOBS) Act in 2012, also known as the Crowdfunding Act, has made possible the provision of securities so that funders, backers, or lenders can earn a positive financial result from their investment (Robertson & Wooster, 2015). Equity crowdfunding is rare, and most rewards take other forms of payment, such as royalty sharing or some form of non-pecuniary rewards (Mollick & Robb, 2016).

Crowdsourcing projects can be analyzed by the level of complexity of the given task, by the relation between the contributions, or by their model of management, as discussed below (see Table 7.1).

Table 7.1 Matrix on crowdsourcing dimensions

Dimensions	Variables		
Level of complexity	High-complexity (project proposal)	Mid-range complexity	Low-complexity (minimal skill)
Relation between contributions	Filtered	Aggregated	
Management of the crowd	Centralized/ decentralized	Open/closed	Collaborative/unknown participation

Source: own elaboration.

Complexity level of the tasks

Crowdsourcing involves the use of crowd labor to perform a specific task, which may require various levels of skill.⁴ At the complex end of the spectrum, firms may want the crowd to generate new specific ideas, which may include answering scientific problems, in exchange for a financial reward. The crowd finances the project proposal and only the winners receive a compensation (Ettliger, 2016). An example is the Innocentive website, a platform where seekers present a problem, evaluate solutions proposed by solvers, and choose the best option that receives an award, often financial, in exchange for the intellectual property of the crowdsourced idea.⁵ Crowdsourcing new ideas is often coordinated through competition. A mid-range skill task can be found in virtual call center services and some citizen science projects, which require some form of accountability by the participating individuals for the system to function. Firms may train contributors or may request a certain level of previous

knowledge from participants. Citizen journalism can be considered another example of a mid-range skill task, particularly valuable to gather information from inaccessible areas (e.g., war zones), though not without concerns on the quality of the contributions (Hossain & Kauranen, 2015).

A crowdsourced task requiring minimal skills is the most prevalent. Micro-tasks are often low-skilled tasks and generally involve repetitive and easy activities, such as transcriptions, annotations, validation, and translations, often to be used for training algorithms (Hossain & Kauranen, 2015). Popularity contests of television shows can also be considered examples of aggregated low-skill contributions (Prpić et al., 2015).

There are crowdsourcing projects where the crowd is in fact not aware of their contribution. This is the case of the CAPTCHA system used to discern humans from computers, giving Google about 83,000 hours of daily labor to support optical character recognition (OCR) and semantic links (for image identification) (Prpić et al., 2015). CAPTCHA stands for a Completely Automated Public Turing test to tell Computers and Humans Apart. It was adopted by Google to digitize books and newspapers, with variations in reCAPTCHA and later including image identification to train algorithm systems.

Aggregate and filtered content

In addition to the level of complexity of a task, and the distribution of the costs, crowdsourcing projects can be organized based on the relation between contributions to achieve the final goal. Some projects break up the process into small parts, such as micro-tasks and voting, which are later aggregated to create a final product (Prpić et al., 2015). A well-known example of this is the crowdsourcing Zooniverse platform,⁶ where the crowd is asked to tag, match, classify, or

identify characteristics of images and videos from large datasets generated by space telescopes or cameras in natural parks. The data collected is then aggregated and analyzed to develop new insights about galaxies or species behavior. The individual contributing members of the crowd do not need to interact, and the task started by one can be continued by another at any time. The platform Zooniverse also provides humanities content such as hand-written texts and asks the crowd to transcribe them, following the same principle.

A similar example is the Foldit project developed by the University of Washington. The crowd plays with protein molecular structures to identify images (e.g., density maps). Results are more accurate interpretations of protein structures than those from automated algorithms, and the crowd can process images faster than previous cryo-electron micrographs or X-ray crystallographs (Bkoep, 2019). The aggregated micro-tasks have served to publish a series of scientific papers by the university team, who kindly acknowledge the contributions of the Foldit players (estimated to surpass the 57,000).⁷

In contrast, crowdsourcing projects can rely on one individual to create a finished product, which may be then selected by the requesting firm and adopted for production. This is the case with many idea generation projects where firms outsource the creation of an innovation or the anticipation of predictions (Hossain & Kauranen, 2015). An example is NASA's \$30,000 award to a solution for the prediction of solar flares (Nam, 2011).

Management of the crowd

Crowdsourcing projects can allow centralized or decentralized authority, open (e.g., interoperability, future reuse of the gathered information) or closed participation (e.g., anonymous participation, peer review), and a range of relationships (from unknown participation

to full collaboration). Activities can be objective, as transcribing a text, or subjective, as voting for “the best of.” Firms may request funds to finance a project, devise a tournament type crowdsourcing (giving a prize for the best idea), call for virtual labor (made up of a series of micro-tasks), or engage in an open collaboration (such as a wiki) (Prpić, Taeihagh & Melton, 2015).

Not all crowdsourcing projects are collaborations between crowdsourcer and crowdsourcer. Projects such as Google’s CAPTCHA tap into the crowd to do a job without encounters or relationships between actors. Platforms such as del.icio.us or YouTube are not clear in their crowdsourcing nature, as the call, the benefits, the tasks, the nature of the participation, and the role of the initiator are not clearly identified (Estelles-Arolas & Gonzalez-Ladrón-De-Guevara, 2012). Such aggregators utilize the crowd’s input, including content, behavior, and ratings, to prioritize content output without calling the crowd for a greater goal. Most crowdsourcing projects, however, do call the crowd.

Attracting the right crowd and aligning efforts to the project’s goal is essential, though not as easy as one would think. Crowds respond to information about the firm or platform where the project takes place as well as to signals from other participating members of the crowd (Belleflamme et al., 2019). Breaking up problems into tasks doable for individuals outside of the firm requires sufficient organizational capacity from the firm in order to return usable results, which explains the emergence of specialized crowdsourcing platforms with a history of constructing crowds, offering paid services to intermediate coordination (Prpić et al., 2015). Consistent and good quality participation from the crowd requires a clear central idea, as well as transparent governance structure. An important point is the ownership of the final product. An example of a jointly owned product is open source software (OSS), free to use and developed by

a crowd generally driven by altruistic motivations, which often outperforms commercial products. Well-known examples include Linux, Mozilla Firefox, OpenOffice, and Ubuntu. OSS is dependent on the active crowd to assess, improve, and evaluate the software to increase its quality, and to keep it usable (Wexler, 2010).

Another example is the free online encyclopedia Wikipedia, with over 49 million articles (6 million of which are in English), edited by 200,000 contributors every month, available in 307 languages, and viewed by over 21 billion individuals across the world on a monthly basis.⁸ The project is one of many products coordinated by the Wikimedia Foundation as part of a free knowledge infrastructure aiming at giving access to all the knowledge of the world.⁹ It has become an important global information resource, a socially constructed common good (Hess, 2012).

Crowdsourcing and digital heritage

Many heritage institutions have a certain experience with crowds providing for finances, information, labor, or ideas in a small scale through the long history of working with volunteers. Adopting a crowdsourcing approach as we now know it involves a larger number of respondents at a much higher speed. For this, constraints to adopt a crowd approach include issues of validating contributions (Ridge, 2013) or the fear of undermining the authority of heritage experts who want to protect a certain exclusivity of the sector (Abbing, 2019). Further, while digital networked technology promises to facilitate referencing by linking to primary data (including heritage collections), the challenges of dealing with the nature of heritage data, characterized by being fuzzy, heterogeneous, uncertain, and interpretative, have not been solved (Blanke et al., 2012).

The issue of multiple objectives may also be identified. While for-profit firms may call the crowd to lower costs of production or reduce risk of product uncertainty as in the case of the creative industries (Dalla Chiesa & Handke, 2020), heritage institutions may seek to engage consumers. Heritage services are often driven by multiple goals, including the preservation and development of cultural heritage, advancing education, providing entertainment, or increasing welfare (Towse, 2019), all of which may be pursued during a crowdsourcing project. Similarly, consumers may want to engage with heritage content for various reasons. Participants in heritage crowdsourcing are motivated to take part in a history- and archaeology-related project; to learn about the subject; to connect with the institution; to develop skills, for fun, out of curiosity; to help out; to contribute to knowledge production; and to learn more about a personal cultural background (related to identity forming) (Bonacchi et al., 2019).

Involving the crowd is inherent to intangible heritage and most forms of culture. In the visual arts, Spencer Turnick can serve as prime example of volunteers contributing as models for a group photograph, including the 18,000 nudes posing at Mexico City's central square.¹⁰ A film example can be found in Peter Jackson's call of 20,000 volunteers to contribute to the sound of the org armies in the *Lord of the Rings: Two Towers* (Prpić et al., 2015). The fashion firm ModCloth has launched several crowdsourcing projects in exchange for discounted purchases, including calling the crowd to contribute images modeling clothes, to suggest clothes based on body types, to vote the best drafts to be produced as final designs, and to provide sketched ideas for new designs, which were in turn voted by the same crowd and adapted into the following season (Estelles-Arolas & Gonzalez-Ladrón-De-Guevara, 2012). Creative industries crowdsourcing regularly involve larger crowds for content generation, ratings, and new ideas,

which do not compare to most heritage crowdsourcing projects, receiving support from a few hundred individuals and rarely surpassing the five thousand (Carletti et al., 2013).

An overview of projects in the heritage field that tap into the crowd using digital technologies has not yet been inventoried and has only recently attracted broad attention (for an overview of a selection of projects from the early 2010s, see Ria, 2014). A few projects with large consortia have been financed by public money in the last decade. Examples include 3D-ICONS, a three-year project for the documentation of archaeological monuments and historic buildings to feed Europeana;¹¹ The Scottish Ten, a project to document five international heritage sites to help with their conservation and management;¹² or the Million Image Database, a repository of heritage images to document the past and to explore the use of 3D imagery.¹³

Crowdsourcing projects specific to advance digital collections of heritage institutions, the GLAMs, have emerged to explore alternative sources of labor, and sometimes funds, for the digitization production process. A recent European survey estimated that museums spend a yearly median €10,000 (mean €111,000) in digital activities and allocate an average of five paid staff and two volunteers toward digital collection activities. Management (e.g., licenses, online services) make up about 40% of the structural costs, digital preservation and archiving make up about 35% of the costs, and the rest is allocated toward user outreach, usage analysis, and editorial work. Eighty-nine percent of the budget is internally generated, and so far, it has been estimated that 31% of museum collections have been digitized (Nauta et al., 2017).

Invariably, availability of online collections increases their use, yet production and proper online publication remain a challenge. To give an example: one large city archive has 32 kilometers of materials, which would represent about 224 million scans (there are roughly 7,000 scans per meter). At an output of 10,000 scans per week, which is quite an impressive mass digitization

process, it would take the archive 431 years to digitize their entire collection. Digitization started in 2001 and continues to improve, speeding up the process to about two million scans per year, or 20,000 scans per week, or 30 scans per minute. This included 6,000 scans requested by 1,260 different consumers (scan on demand program). Since 2001, the archive has digitized six million scans (Holtman & Van Zeeland, 2019). However, digitized material can only be accessed when its contents can be searched for and identified; a digital image in itself is not enough. The city archive decided to call the crowd to support the identification of important information, such as dates, names, places, and type of document. With a grant of €150,000 from the Mondriaan Foundation, the Amsterdam City Archive collaborated with a software partner, Picturae, to develop the crowdsourcing software, which is currently used by many more heritage institutions to annotate digital collections. To ensure quality, every document data element is entered twice (SAA, 2013). The archive commissioned the creation of an online program called Many Hands (Vele Handen) to coordinate crowdsourced annotations and in the first 1.5 years, over 1.6 million objects were identified with name, birth place, and date of birth twice, representing a full-time staff working for a period of over 18 years (SAA, 2013).¹⁴ While the onsite access to the archives has decreased by 66% in the past ten years, online access continues to increase (from zero to 18,000 yearly requests in the same period) (Duran, 2019).

Besides tapping into the crowd to advance the digitization of collections by providing financial resources or labor to enhance existing collections, crowdsourcing projects can support public engagement and can lead to developing new collections: for instance, to document history as a form of intangible heritage (Carletti et al., 2013). The main crowdsourcing forms in the heritage field are crowdfunding (acquiring funds to finance a project), idea generation projects (for new products and services), aggregated input projects (generally to speed digitization projects, some

of which through micro-tasks), and open collaboration projects (for instance to populate a digital collection). We follow with a discussion of each with examples (summarized in Table 7.2).

Table 7.2 Typology of crowdsourcing in heritage projects

Type of project	Examples		
Crowdfunding projects	Hardware/software development	Production of digital surrogates	Sustainability (none found)
Idea generation projects	Reuse of collection	For product development (shop)	
Aggregated input projects	Classification, transcription, correction, revision	Curation, preservation	Ownership identification (orphan works)
Open collaboration projects	Development of a common resource	Individual/institutional	Use of own/third-party platform

Source: own elaboration.

Crowdf unding projects

While heritage institutions increasingly engage in development campaigns, asking private patrons for financial support, projects inviting the crowd to co-finance digitization projects or the

production of digital surrogates are extremely rare. Generally, crowds are asked to support the financing of capital projects, the acquisition, preservation, or exhibitions of collection objects, and the financing of events in museums but rarely a digital collection. A well-known portal for arts and culture crowdfunding is Kickstarter.

An example of co-financing hardware is the crowdfunding project to acquire 35 Oculus Rift 2 glasses to present a 3D video of the city of Horn in the 1650s. The Westfries museum developed a video and wanted to show it using advanced technology, calling the crowd through the Dutch cultural crowdfunding portal Voordekunst, receiving 112% of its €32,000 goal, backed by 317 donors.¹⁵

A notable crowdsourcing project for a digital collection is the call by the information and library services department of the ethnographic museum in Amsterdam. The museum called the crowd to raise €10,000 for the development of a search engine to access 355 photo albums and to develop a mobile phone application for the documentation of the collection. The museum received 13 crates with photo albums left in Java and Sumatra when Dutch-Indians moved to the Netherlands after the war. The majority found their owner family, but 355 personal photo albums remained unidentified. The museum developed a “Photo seeks family” project online to involve younger generations and to reach non-visitors who were asked to identify their family members’ stories (KIT, 2013). The museum used Voordekunst and received 125% of the amount requested, financed by 232 crowd donors.¹⁶

A similar project was developed by the Amsterdam City Archive where the crowd was invited to co-finance the digitization of 38,463 photographs, including glass negatives. The project was launched in 2013 and, at the time of writing, it has digitized 12,000 photographs and received €34,500 of a goal of €50,000. The project has its own website, where it was possible to donate

and where a special page acknowledges donors. The crowd is also invited to provide additional information to identify the photographs: for example, the individuals portrayed, the locations, or the dates. All the digitized collection is still available online.¹⁷

Idea generation projects

The use of digital collections for idea generation is not always crowdsourced, as online consumers are free to make use of the images for any purpose without collaborating with the museum institution. The Rijksmuseum in Amsterdam has devised a crowdsourcing system for ideas on the reuse of collections through the bi-yearly Rijksstudio Award. The Rijksstudio is the portal of the Rijksmuseum online collection, which is the base for the idea generation. The 2020 prize amounts to €7,500, a Young Talent Award of €5,000 will be awarded, and the winner of the public vote will receive €25,000. At the time of writing, entries were still being evaluated by a jury of experts and the crowd was called to vote for a favorite. In addition to the prize, winners can become part of the goods sold at the museum's gift shop.¹⁸

Aggregated input projects

Most crowdsourcing projects ask the crowd to contribute to the classification, transcription, correction, revision, and documentation of collections to enrich metadata and contextualization. Perhaps the first effort to improve access to digitized collections through crowdsourcing is the Steve.museum, a collaboration initiated by the Metropolitan Museum of Art, the Guggenheim Museum, the Cleveland Museum of Art, and the Deventer Art Museum. It was launched in 2005 when museums were just broadly adopting the web to present collections online. The Steve.museum project asked the crowd to support classification of collections through tagging

objects and in doing so develop a folksonomy. From the initial 30 objects that were given crowd terms, 80% of these were not found in the museum's documentation system but added on. The project required a new data model infrastructure to capture the crowd's contributions (Chun et al., 2006).¹⁹ A more technically advanced semantic annotation project was Accurator, led by the Rijksmuseum, designed as niche-source to provide expert knowledge by the crowd. The bird experts identified 1,528 birds in 213 artworks in the collection (Jongma & Dijkshoorn, 2016).²⁰ The Accurator project was a collaboration with a consortium of Dutch researchers from the VU University, Delft University of Technology, Centre for Mathematics and Informatics, and three large museums—Rijksmuseum, Naturalis, and Sound and Vision (Jongma & Dijkshoorn, 2016). These are examples of innovative explorations to involve the crowd in collection annotation. Crowdsourced tagging can serve other purposes. The Indianapolis Museum of Art aggregated tags into heatmaps to identify the most visible areas of an artwork (Carletti et al., 2013). Such a by-product of an initial tagging project may become more prevalent as the availability of data allows for new insights in consumer behavior online.

Crowdsourcing the curation of digital collections has proven quite successful to engage with the public. The crowd can support the selection, classification, and organization of an existing collection (Carletti et al., 2013). These can blend the digital and physical worlds, such as the Dutch Kröller-Müller museum project, first for adults and then for children, asking the crowd to rank works. Those receiving greater votes were included in a physical exhibition (Schoemaker, 2010). Similarly, the Victoria and Albert Museum devised a voting system to rank its 140,000 digital images of the collection to appear in the Search the Collections page.²¹

A geographic-based crowdsourcing project of sounds was led by the British Library Sounds collection, which presented over 50,000 recordings from the library collection and invited the

public to contribute their own sounds. A series of Sound Maps were created documenting accents (the crowd could read a children's story or a list of six words) and daily sounds at home, work, or play (over 2,000 recordings were received by 350 contributors).²² A similar project collected 1001 Stories of Denmark, in a crowdsourcing effort initiated by the government meant to capture intangible heritage through geo tagging. At the time of writing, the project website was no longer functional, though recorded stories were accessible via the project's YouTube channel.²³

A more challenging effort was led by EnDOW (Enhancing access to 20th-century cultural heritage through Distributed Orphan Works clearance). The crowd was called to support rights clearance of European digitized cultural heritage by identifying orphan works. The project was a response of the European Directive on orphan works that required institutions to perform diligent search before publishing collections online. The labor costs that such diligent search would require are prohibitively expensive for any cultural institution with a sizable collection. The crowd is instructed on how to search for a work's potential right holder before labeling it as orphan (Borghi, Erickson & Favale, 2016).²⁴

Open collaboration projects

There are crowdsourcing projects that invite the creation of a common resource. These tend to be larger projects with a strong crowd engagement, for long periods of time. One Dutch example is the whale stranding database,²⁵ launched in 2006 by the Naturalis Biodiversity Centre to merge the historic data housed at the museum, including the first documented case dating from 1,255, and inviting the crowd to add current findings. The site provides basic fields to harmonize the documentation of the findings, including date, location, measurements, and an image when

available. The site also provides basic clues to identify the various characteristics of the animals, to differentiate a whale from a dolphin for example (Naturalis, 2006). The site is currently an important resource where over 12,000 whale and dolphin strandings have been documented (513 entries in 2019), available for everybody to use free of charge. The natural history museum has in fact taken the role of coordinator of the newly formed common resource.

The Amsterdam museum intends to similarly serve as facilitator of a socially constructed commons of intangible heritage about the city of Amsterdam, starting with the Memory of East project in 2006, followed by Centrum and West, and eventually to expand to South, South East and North neighborhoods of the city. The overarching Memory of Amsterdam project intends to collect, preserve, and make available the daily memories and stories of residents of the city. Selected content is regularly used for thematic exhibitions (AM, 2018). The crowd is similarly guided on how to generate content through a series of harmonizing fields, such as place, date, and keywords.²⁶

An institutional crowdsourcing project of heritage content across Europe is Europeana, currently giving access to 50 million items. Though the project receives a yearly subsidy from the European Commission (over €6 million in 2018), this is largely allocated to cover staff and general operations (80% of the budget), as well as the development of the portal and regular meetings to coordinate the project (20%) (Europeana, 2018). Participating heritage institutions bear the costs of digitization of collections and ingestion into the national aggregator. Institutions receiving European grants are required to publish the digitized outcomes in the Europeana repository. The portal launched in 2007 and has evolved from earlier efforts to integrate heritage collections remotely across Europe (Navarrete, 2014).

Whereas the previous examples have a coordinating institution directing the crowdsourcing efforts, there are other examples where a platform is adopted as common storage of content, without clear curatorial guidelines. Such is the example of the Sketchfab platform, launched in 2011 for the publication of 3D, VR, and AR content. Museums can upload content, some have created a profile, though the majority of the content is uploaded by member of the crowd.²⁷ This portal gives access to the largest 3D, VR, and AR collection of heritage in the world (consumers do have to filter through all the non-heritage-related content!).

Conclusion and managerial implications

Digital technology has transformed the social perception of crowds as “collective intelligence that solves problems” (Wexler, 2010, p. 12) voluntarily, with important economic implications. Tapping into the crowd as a pool source of labor may crowd out payment for activities to trained individuals in an already distressed labor market; 66% of Dutch museum staff is currently non-paid volunteer or intern (NMV, 2018). In addition, a study of Swiss museums reported that involving the crowd to support digitization of collections was clearly seen to deliver greater risks than benefits (Estermann, 2013). On the other hand, there are four important positive outcomes. First, the crowd can advance work that has been limited by the available resources to heritage institution. Particularly, regarding the digitization of collections, crowds can advance the identification of objects so that these are usable for any purpose. From one account, the crowd advanced in one and a half year what a full-time staff would have taken 18 years to do. Second, the engagement with the institution and the content can lead to new forms of citizen participation: for instance, to create socially constructed commons. From the presented examples it can be argued that the digital repositories, freely accessible to all, are a valuable resource in the

information economy. Individuals from every field, as well as machines, can benefit from quality, structured, diverse, reliable, and sustainable content.

As with any resource, a digital repository requires investment and maintenance for its sustainability (and growth). Conceptualizing the digital heritage repository as a socially constructed cultural commons may strengthen engagement with communities of users to sustain it. Stimulating reuse of objects in the digital heritage repository can be expected to support sustainability of the increasingly dynamic repositories, able to adapt to social changes and crowd contributions.

Third, crowdsourcing may improve the distribution of costs and benefits beyond political borders. Currently, many digital heritage projects are paid for by tax payers, yet online dissemination allows for consumption beyond the paying population. Crowdsourcing could provide input from a greater segment of the population, including those outside of the political district.

Crowdsourcing and niche-sourcing may favor popularity as network effects inevitably have strong dynamics. It is the role of policy makers to intervene and support greater diversity and equity in the access and development of heritage. Because heritage is socially defined and changes over time, it can be expected that priorities will change.

Last, engaging with crowdsourcing and digital heritage content would inevitably lead to a digital trail that may reveal trends in emerging forms of digital cultural consumption. So far, much of the digital content has been analyzed by private firms (such as the family of Google services), yet little knowledge has been shared with researchers, politicians, and the crowds. The explosion of content, particularly through social media, is simply not sustainable. Curating a socially

constructed heritage commons may raise awareness of the terabytes needed to maintain our current digital consumption lifestyle.

References

- Abbing, H. (2019). *The changing social economy of art: Are the arts becoming less exclusive?* Cham: Palgrave Macmillan.
- Aghaei, S. (2012). Evolution of the World Wide Web: From Web 1.0 to Web 4.0. *International Journal of Web & Semantic Technology*, 3(1), 1–10.
- Amsterdam Museum. (2018). *Amsterdam Museum jaarverslag 2018*. Amsterdam: Amsterdam Museum. Available at <https://www.amsterdammuseum.nl/over-ons/over-de-organisatie/jaarverslagen>.
- Bell, A. E. (1947). *Christian Huygens and the development of science in the seventeenth century*. *Journal of the Franklin Institute*, 246(3), 264–264. [https://doi.org.eur.idm.oclc.org/10.1016/0016-0032\(48\)90085-4](https://doi.org.eur.idm.oclc.org/10.1016/0016-0032(48)90085-4).
- Belleflamme, P., Lambert, T., & Schwienbacher, A. (2019, August 9). *Crowdfunding dynamics*. Social Science Research Network. <http://dx.doi.org/10.2139/ssrn.3259191>.
- Belleflamme, P., & Peitz, M. (2020). Ratings, reviews and recommendations. In R. Towse & T. Navarrete (Eds.), *A handbook of cultural economics* (3rd ed., pp. 466–474). Cheltenham: Edward Elgar.
- Berg, J., Furrer, M., Harmon, E., Rani, U., & Six Silberman, M. (2018). *Digital labour platforms and the future of work: Towards decent work in the online world*. International Labour Office. https://www.ilo.org/wcmsp5/groups/public/---dgreports/---dcomm/---publ/documents/publication/wcms_645337.pdf
- Berners-Lee, T., Cailliau, R., Groff, J.-F., & Pollermann, B. (1992). World-Wide Web: The information universe. *Electronic Networking: Research, Applications and Policy*, 2(1), 52–58.

- Bertacchini, G., Marrelli, M., & Santagata, W. (2012). Defining cultural commons. In G. Bertacchini, M. Marrelli, & W. Santagata (Eds.), *Cultural commons* (pp. 3–18). Cheltenham: Edward Elgar.
- bkoep. (2019, November 11). *The Foldit cryo-EM paper*. foldit.
<https://fold.it/portal/node/2008272>
- Blanke, T., Bodard, G., Bryant, M., Dunn, S., Hedges, M., Jackson, M., & Scott, D. (2012). Linked data for humanities research—The SPQR experiment. In *2012 6th IEEE international conference on digital ecosystems and technologies (DEST)*. IEEE.
<https://ieeexplore.ieee.org/document/6227932>
- Bonacchi, C., Bevan, A., Keinan-Schoonbaert, A., Pett, D., & Wexler, J. (2019). Participation in heritage crowdsourcing. *Museum Management and Curatorship*, *34*(2), 166–182.
- Bonancini, E. (2013). Stories on geographies: Geo-social tagging for co-creation of cultural value. *International Journal of Heritage in the Digital Era*, *2*(2), 222–243.
- Borghi, M., Erickson, K., & Favale, M. (2016). With enough eyeballs all searches are diligent: Mobilizing the crowd in copyright clearance for mass digitisation. *Chicago-Kent Journal of Intellectual Property*, *16*(1), 135–164.
- Carletti, L., McAuley, D., Price, D., Giannachi, G., & Benford, S. (2013). Digital humanities and crowdsourcing: An exploration. In N. Proctor & R. Cherry (Eds.), *MW2013: Museums and the Web 2013* (n.p.). Museums and the Web.
<https://mw2013.museumsandtheweb.com/paper/digital-humanities-and-crowdsourcing-an-exploration-4/>
- Cherry, M. (2011). A taxonomy of virtual work. *Georgia Law Review*, *45*(4), 951–991.

- Chun, S., Cherry, R., Hiwiler, D., Trant, J., & Wyman, B. (2006). Steve.museum: An ongoing experiment in social tagging, folksonomy, and museums. In J. Trant & D. Bearman (Eds.), *Museums and the web 2006: Proceedings* (n.p.). Archives & Museum Informatics.
<https://www.museumsandtheweb.com/mw2006/papers/wyman/wyman.html>
- Dalla Chiesa, C., & Handke, C. (2020) Crowdfunding. In R. Towse & T. Navarrete (Eds.), *Handbook of cultural economics* (3rd ed., pp. 158–167). Cheltenham: Edward Elgar.
- Duran, C. (2019). *Analysis of the use of the collections at the Amsterdam City archives (1999–2018)*. Amsterdam: Stadsarchief Amsterdam. Internal report.
- Estellés-Arolas, E., & González-Ladrón-De-Guevara, F. (2012). Towards an integrated crowdsourcing definition. *Journal of Information Science*, 38(2), 189–200.
- Estermann, B. (2014). Diffusion of open data and crowdsourcing among heritage institutions: Results of a pilot survey in Switzerland. *Journal of Theoretical and Applied Electronic Commerce Research*, 9(3), 15–31.
- Ettliger, N. (2016). The governance of crowdsourcing: Rationalities of the new exploitation. *Environment and Planning A: Economy and Space*, 48(11), 2162–2180.
- European Commission. (2011). *The new renaissance: Report of the 'Comité des Sages'*.
https://ec.europa.eu/commission/presscorner/detail/en/IP_11_17
- Europeana. (2017). *Europeana strategy 2020*. <https://pro.europeana.eu/post/europeana-strategy>
- Europeana. (2018). *Annual report: A decade of democratising culture*. Published online 29 March 2019 at <https://pro.europeana.eu/page/annual-report-2018-a-decade-of-democratising-culture>.

- Gellers, J. C. (2016). Crowdsourcing global governance: Sustainable development goals, civil society, and the pursuit of democratic legitimacy. *International Environmental Agreements: Politics, Law and Economics*, 16(3), 415–432. <https://doi.org/10.1007/s10784-016-9322-0>
- Hess, C. (2012). Constructing a new research agenda for Cultural Commons. In G. Bertacchini, M. Marrelli, & W. Santagata (Eds.), *Cultural commons* (pp. 19–35). Cheltenham: Edward Elgar.
- Holtman, Marc and Van Zeeland, Nelleke (2019) *Het jaar van 6 miljoen. Jaarverslag digitalisering 2018 Stadsarchief Amsterdam*. Amsterdam: Gemeente Amsterdam Stadsarchief.
- Hossain, M., & Kauranen, I. (2015). Crowdsourcing: A comprehensive literature review. *Strategic Outsourcing: An International Journal*, 8(1), 2–22.
- Howe, J. (2008). *Crowdsourcing: Why the power of the crowd is driving the future of business*. New York: Three Rivers Press.
- International Telecommunications Union. (2018). *Measuring the information society report* (Vol. 2). <https://www.itu.int/en/ITU-D/Statistics/Documents/publications/misr2018/MISR-2018-Vol-2-E.pdf>
- Jongma, L., & Dijkshoorn, C. (2016). Accurator: Enriching collections with expert knowledge from the crowd. In D. Bearman & J. Trant (Eds.), *MW2016: Museums and the web 2016* (n.p.). Archives & Museum Informatics. <https://www.museumsandtheweb.com/mw2006/papers/wyman/wyman.html>
- Koninklijk Instituut voor de Tropen. (2013). *KIT annual report 2012*. Amsterdam: KIT. Available at <https://www.kit.nl/about-us/annual-report/>.
- Kushner, S. (2016). Read only: The persistence of lurking in Web 2.0. *First Monday*, 21(6), n.p.

- Mollick, E., & Robb, A. (2016). Democratizing innovation and capital access: The role of crowdfunding. *California Management Review*, 58(2), 72–87.
- Naturalis. (2006). *Annual report 2006*. Leiden: Naturalis Biodiversity Centre.
- Nauta, G.-J., van den Heuvel, W., & Teunisse, S. (2017). *Europeana DSI 2—Access to digital resources of European heritage. D4.4. Report on ENUMERATE Core Survey 4*. Digitaal Erfgoed Nederland.
https://pro.europeana.eu/files/Europeana_Professional/Projects/Project_list/ENUMERATE/deliverables/DSI2_Deliverable%20D4.4_Europeana_Report%20on%20ENUMERATE%20Core%20Survey%204.pdf
- Navarrete, T. (2014). *A history of digitisation: Dutch museums*. (Doctoral dissertation, University of Amsterdam). <https://hdl.handle.net/11245/1.433221>
- Nederlandse Museumvereniging. (2018). *Museumcijfers 2018*.
https://www.museumvereniging.nl/media/museumcijfers_2018_def.pdf
- Prpić, J., Shukla, P. P., Kietzmann, J. H., & McCarthy, I. P. (2015). How to work a crowd: Developing crowd capital through crowdsourcing. *Business Horizons*, 58(1), 77–85.
- Prpić, J., Taeihagh, A., & Melton, J. (2015). The fundamentals of policy crowdsourcing. *Policy & Internet*, 7(3), 340–361.
- Ranard, B. L., Ha, Y. P., Meisel, Z. F., Asch, D. A., Hill, S. S., Becker, L. B., Seymour, A. K., & Merchant, R. M. (2014). Crowdsourcing—Harnessing the masses to advance health and medicine, a systematic review. *Journal of General Internal Medicine*, 29(1), 187–203.
- Ridge, M. (2013). From tagging to theorizing: Deepening engagement with cultural heritage through crowdsourcing. *Curator: The Museum Journal*, 56(4), 435–450.

Robertson, E., & Wooster, R. B. (2015, July 15). *Crowdfunding as a social movement: The determinants of success in Kickstarter campaigns*. Social Science Research Network.

<https://doi.org/10.2139/ssrn.2631320>

Schoemaker, T. (2010, November). Kinderen kiezen kunst in museum Kröller-Müller. *NRC Handelsblad*. <https://www.nrc.nl/nieuws/2010/11/26/kinderen-kiezen-kunst-in-museum-kroller-muller-11974501-a79233>

Stadsarchief Amsterdam. (2013). *Open Archieven door Crowdsourcing: een project voor grootschalige archiefontsluiting*. Eindverslag voor het Mondriaan Fonds. Amsterdam: Stadsarchief Amsterdam. Internal report.

Towse, R. (2019). *A textbook of cultural economics*. Cambridge: Cambridge University Press.

Visser, J. (2014). Strategies for a heritage revival in the digital age. In L. Egberts & K. Bosma (Eds.). *Companion to European heritage revivals* (pp. 73–93). Cham: Springer.

Wexler, M. N. (2011). Reconfiguring the sociology of the crowd: Exploring crowdsourcing. *International Journal of Sociology and Social Policy*, 31(1/2), 6–20.

¹ The Web 1.0 is characterized by static document publication which evolved into the Web 2.0 when it allowed user contributions (such as likes, uploading of media, and writing content) and interactions (such as discussions and mashups). The Web 3.0 refers to the Semantic Web, where content is encoded to allow machines to understand it (such as a name, a date, or a place). Some envision the Web 4.0, or Symbiotic Web, to be characterized by a seamless relation between machines and humans through robotics (such as virtual assistants) and high personalization (such as context specific displays) (Aghaei et al., 2012).

² The original call via Twitter is available at <https://twitter.com/LuisCollantesR/status/1143170732450570240>. The information about the painting with mystery location is available at <https://www.museodelprado.es/coleccion/obra-de-arte/paisaje/07e8f75e-3622-4b48-ba3d-2fc1096e6c5a?searchMeta=agustin%20riancho>.

³ Data from the project is available at <http://data.myworld2015.org/>.

⁴ For an overview of large crowdsourcing projects with a short description see Cherry (2011).

⁵ <https://www.innocentive.com/>

⁶ <https://en.wikipedia.org/wiki/Zooniverse>.

⁷ <https://fold.it/portal/info/about>.

⁸ All statistics from Wikipedia are available at <https://en.wikipedia.org/wiki/Wikipedia:Statistics>.

⁹ <https://wikimediafoundation.org/>.

¹⁰ <https://www.spencertunick.com/>.

¹¹ <http://3dicons-project.eu/>.

¹² <https://www.engineshed.scot/about-us/the-scottish-ten/>.

¹³ <https://www.millionimage.org.uk/>.

¹⁴ The projects can be found at <https://velehanden.nl/>.

¹⁵ A portion of the video is available at <https://vimeo.com/105549924>. The Crowdsourcing page is available at <https://www.voordekunst.nl/projecten/2751-de-gouden-eeuw-vr-stap-in-1650>.

¹⁶ The successfully closed crowdsourcing page is available at <https://www.voordekunst.nl/projecten/578-foto-zoekt-familie-1>.

¹⁷ <http://redeenportret.nl/doneer>.

¹⁸ The Rijksstudio is available at <https://www.rijksmuseum.nl/en/rijksstudio>.

¹⁹ A folksonomy was desired to allow non-experts to search and find objects in collections that would otherwise be hard to find in the authored expert information system. The name clearly differentiates the folk nature of the taxonomic classification. Terms were often simple descriptors, such as animal, archery, baroque, luxury, movement, nude, ornate, saddle, and woman to describe a 1620 silver “Diana and the Stag” by Joachim Friess, part of the Metropolitan Museum of Art (Chun et al., 2006).

²⁰ The project website is available at <http://accurator.nl>.

²¹ The dynamic web page displays a new set of objects based on the crowd’s input (<https://collections.vam.ac.uk/>).

²² The crowdsourcing project took place in 2010, which is available together with the rest of the sound collection (<https://sounds.bl.uk/Sound-Maps>).

²³ For a short overview of the project see <https://www.youtube.com/watch?v=NFw2x20y6vo>.

²⁴ The project website is available at <http://diligentsearch.eu/>.

²⁵ The project can be found at <http://www.walvisstrandingen.nl/>.

²⁶ The project is available at <https://geheugenvan.amsterdam/>.

²⁷ The British Museum, for example, opened a profile in 2014 and currently displays 257 models (see <https://sketchfab.com/britishmuseum>).