Field Notes: An Exploration of Crowdsourcing Platforms for Natural History Collections

Digital Curation Research Final Research Paper

Prepared by Lauren Cohen

laurencohenpgh@gmail.com

AS.460.674.81.SU21: Digital Curation Research Paper

Museum studies Digital Curation Certificate Program

Johns Hopkins University

August 14, 2021

ABSTRACT

Drawing from a long tradition of public participation and recent innovations in digital technology, natural history museums (NHMs) are transforming the way they conduct research and engage their audience. Digitization is unlocking a wealth of biodiversity data and expanding the use of natural history collections (NHCs) across the globe. Online citizen-science crowdsourcing is imploring the aid of a massive network of amateur-experts to assist with the transcription of historical records which cannot be completed by computers. Crowdsourcing platforms such as Zooniverse, DigiVol, and the Smithsonian Transcription Center are attracting a new generation of citizen-scientists and increasing the rate NHMs are able to generate and process biodiversity data. This symbiotic relationship is leading to digitized collections becoming more accessible to researchers and advancing our understanding of anthropogenic changes taking place today; expanding the user's appreciation of science by allowing them to actively participate in the discovery process; and resulting in high-quality data for peer-reviewed publications. This research project will explore citizen-science crowdsourcing from its rich history through current use; pinpointing the best online platforms to aid with natural history collection transcription; while also revealing the benefits and concerns associated with amateur-expert powered projects.

TABLE OF CONTENTS

Page

LIST OF FIGURES

Page

INTRODUCTION

Purpose of Research

Digital technology has altered the way in which natural history museums (NHMs) conduct research and interact with their audience. Over the past two decades, NHMs have undergone rapid change and begun to adopt new methods for generating and processing large amounts of biodiversity data. Digitization, in combination with citizen-science crowdsourcing, are the two mobilizing forces in the global effort to improve data discovery and enhance audience engagement.

Online citizen-science crowdsourcing has been identified as a potential tool to aid with the transcription of specimen labels, field notebooks, and correspondences. These biocollections contain abundant historical records and can contribute to the study of anthropogenic changes occurring today (Hill, 2012, p. 220). Crowdsourcing is more than just an instrument for creating and delivering content (Ridge, 2017, p. 2). This researcher agrees that citizen-science generated online transcriptions are helpful because they make the digitized artifacts more accessible online; but crowdsourcing goes beyond this content-creation framework by offering an opportunity for users to do something more than simply consume information (Owens, 2014, p. 279).

Citizen-science crowdsourcing is a relatively new digital model for NHMs; but in terms of collaboration between amateur-expert naturalists, the public's participation in scientific research is not a modern invention (Sforzi, 2018, p. 429). It predates both the professionalization of science, as well as textbook definitions for 'citizen science' and crowdsourcing' (2018, p. 430). Centuries ago, the majority of all scientific study, in North America and Europe, was conducted by non-scientist community members (Miller-Rushing, 2020, p. 17). Many of these amateur-experts carried out research that was indistinguishable, and often superior to the work done by professionals of the time (2020, p. 17). Carl

Linnaeus and Charles Darwin relied heavily on these extensive communities of citizen-scientists to further their hypothesis.

Online citizen-science crowdsourcing platforms, such as Zooniverse, encourage the same two-way knowledge sharing of the past; and the development of web-based citizen science endeavors is forged from a long tradition of utilizing the general public, who have a strong interest in the subject matter, to process data (Hill, 2012, p. 228). From humble beginnings, Zooniverse evolved to become the world's leading citizen science platform with over 2-million registered volunteers (Zooniverse, 2021). Not long after its inception, and overwhelming popularity, other people-powered science projects began to rapidly come online. DigiVol and the Smithsonian Transcription Center are two of the other better-known platforms for transcribing natural history collections. All three of the platforms utilize the internet to disseminate and accelerate biodiversity data capture at extraordinary rates (Hill, 2012, p. 228).

At the present, there are still a substantial number of NHM biodiversity specimens (as many as 1-billion) stored away in cabinets (Ellwood, 2015, p. 387). In the United States, only 10% of these natural history collections (NHCs) are accessible (2015, p. 383). One ambition of the global biodiversity community is to unlock this data for researchers through broadscale digitization; but meeting this ambitious objective requires modernized technology and innovative engagement strategies that extend beyond the physical boundaries of NHMs (2015, p. 383). Citizen-science crowdsourcing is a sustainable model for institutions to consider in order to generate transcriptions for collections, and ease the workload associated with the digitization process. Museums are already grappling with budget and staffing cuts; and it is well known that the transcription process associated with these collections is "time consuming, intellectually intensive, and expensive for an organization to facilitate" (Mika, 2017, p. 61). Crowdsourcing, enabled by the aid of digital citizen-scientists, is the solution.

Figure 1. Biological collections housed at the National Museum of Natural History. [Photo courtesy of Journal of Systematics and Evolution, 2018].

Numerous publications have been released surveying the landscape of current platforms and the digital citizen-science crowdsourcing movement; but in order to fully understand citizen-science crowdsourcing, a hands-on investigation in addition to a literature review adds to this research and expands on the current effects of these digital citizen-science methods on the field. This paper explores the aforementioned citizen-science crowdsourcing platforms in an attempt to comprehend the progress that has been made over the course of the last two decades. By adding first hand examination case studies to this research, key insights can reveal what the current effects are and illuminate its future potential.

Problem Statement

This research project will examine the new dawn of citizen science; and how crowdsourcing can be used as a sustainable model by natural history museums (NHMs) looking to generate transcriptions for large or diverse natural history collections (NHCs). To determine what are the best citizen-science crowdsourcing platforms on the market, the researcher will actively engage and assess three of the most popular- Zooniverse, DigiVol, and the Smithsonian Transcription Center. What makes these three platforms the best options for NHC transcription?

In order to better understand the benefits and concerns associated with citizen-science crowdsourcing, the researcher will conduct a thorough literature review of all relevant published findings to best formulate a response. What are the benefits and challenges associated with utilizing citizen-science crowdsourcing platforms?

In light of all this enquiry, the researcher will seek a definitive answer as to why NHMs are not using citizen-science crowdsourcing more often to transcribe their NHCs. What is causing NHMs not to make use of this method for data discovery and user engagement?

METHODOLOGY

This research project's overall goal was to improve the understanding of available citizen science crowdsourcing platforms that natural history museums (NHMs) can engage with. While also identifying the benefits and challenges associated with the crowdsourcing of natural history collections (NHCs). After careful consideration, the following methodologies were used: a thematic review of relevant literature and three case studies of specific citizen science crowdsourcing platforms.

Thematic review of relevant literature: This research paper follows the traditional structure of a thematic literature review; focusing on one major topic, the citizen science crowdsourcing of natural history collections. Phrases such as "citizen science," "crowdsourcing," and "natural history" were pursued through a variety of bibliographic databases, including JSTOR and online search-engines. A total of 9 peer-reviewed publications and articles were identified and investigated; each offering unique perspectives into the history, available platforms, benefits and concerns associated with citizen-science crowdsourcing.

Case studies of specific citizen science crowdsourcing platforms: Case studies have become an increasingly prevalent research method across various disciplines. This approach focuses on data collection from an array of reputable sources and offers a deeper understanding of the subject matter. By becoming an active transcriber, the researcher was able to gain first-hand experience and record observations from three reputable citizen-science crowdsourcing platforms- Zooniverse, DigiVol, and the Smithsonian Transcription Center. These three platforms all operate fundamentally in similar ways, and are the most popular for transcribing natural history collections. They've also earned global recognition for their large volunteer communities, diverse projects, and published analytical research. This investigative study will focus on those areas as well as each platform's history, website design, volunteer statistics, transcription processes, and engagement practices that foster repeat visits.

LITERATURE REVIEW

History of Citizen Science

It is easy to presume that citizen-science crowdsourcing is a fairly new phenomenon generated by the rise of digital technologies and connected mobile devices (Heckler, 2018, p. 2). Especially at a time when people are so easily linked with one another and have a growing interest in a broad-range of environmental issues. Another reason for this misconception, is that there has been remarkably little written about the history of citizen-science despite its monumental contributions to the field (Miller-Rushing, 2020, p. 22). To be clear, citizen-science and crowdsourcing are neither new or unique data collecting methods. Both have been around as popular practices for centuries, embedded in the broader relational history between science and society (Mahr, 2018, p. 103). From this perspective it is best to begin our discussion of citizen-science crowdsourcing by tracing its history, to explore its impactful contributions and historic roots across a range of genres including biology and ecology.

Research conducted by large networks of amateur-expert naturalists pre-dates the professionalization of science (Sforzi, 2018, p. 430). In the past, a large portion of natural history was powered by citizen-scientists who conducted investigations independently of professionals (Miller-Rushing, 2020, p. 17). The broad field, and history, of citizen-science includes many familiar stories and well-known names. Their collaborative efforts were intended to "unfold the book of nature" through the accumulation and analysis of biodiversity data (Mahr, 2018, p. 104). In the eighteenth century, a Swedish naturalist by the name of Carl Linnaeus was imploring a consortium of amateurs to collect and observe specimens for his binomial naming system (Miller-Rushing, 2020, p. 18). Linnaeus produced instructions on how to properly collect natural history specimens and record their data (Burnett, 2021). His students, known as the Apostles of Linnaeus, embraced his ideas and eventually disseminated the instructions internationally (2021). Linnaeus' instructions would endure for years to come and played a major role in the Enlightenment period's attempt to document the natural world (2021).

By the nineteenth century, Charles Darwin was in the midst of a massive letter-writing campaign with amateur-expert naturalists that would last for decades (Zimmer, 2011). Although the professionalization of science had already taken hold, academics and specialists continued to view the work being conducted by citizen-scientists to be highly valuable (Mahr, 2018, p. 104). Darwin used these correspondences to gather "biological intelligence," to later validate his arguments regarding evolutionary theory (Zimmer, 2011). "In letters he wrote to naturalists around the world, Darwin asked for details about all manner of natural history, from the color of horses in Jamaica to the blush that shame brought to people's cheeks" (2011). Darwin exhausted the nineteenth century postal system, writing upwards of 1,500 letters per year (2011). His network included both men and women, from a diverse spectrum of backgrounds with one common interest- science.

Darwin and Linnaeus, with the aid of citizen-scientists, built some of the most valuable natural history collections (NHCs) in the world. These historical biodiversity samplings- ranging from specimens to records- have been used to develop our modern taxonomic naming system and understand the dynamic forces associated with evolution (Miller-Rushing, 2020, p. 19). Throughout the twentieth century fewer and fewer of these citizen-science projects occurred. In 1900, American Museum of Natural History ornithologist, Frank Chapman, launched the Christmas Bird Count (Sforzi, 2018, p. 430). It is the longest-running active citizen-science project on record (2018, p. 430). Instead of hunting the birds, which was a holiday tradition, Chapman proposed a count to aid in the understanding of declining populations. This ongoing citizen-science project, overseen now by the National Audubon Society, has contributed to over 200 publications which have relied heavily on the resulting data (Kosmala, 2016, p. 551).

Since then, the role of amateur-experts as colleagues to those professional scientists has lessened (Miller-Rushing, 2020, p. 18). Despite the change in the way research is being conducted presently there

is a growing number of naturalist clubs, institutions, and citizen-science crowdsourcing platforms left to salvage this once flourishing enterprise (2020, p. 18).

Natural History Museums and Digitization

Since their establishment, natural history museums (NHMs) have continued to work closely with amateur-experts to build a greater understanding of the natural world (Sforzi, 2018, p. 430). It is through this collaboration that many natural history collections (NHCs) were developed and maintained. Over the past few decades, NHMs have reimagined their relationship with the citizen-science community and revitalized their function within society (2018, p. 429). Alongside the traditional role of a treasure house, NHMs have begun to actively engage a broader audience through the digitization of their collections.

With the advent of the internet and rise in new technology, NHCs are becoming more accessible to new generation citizen-scientists and researchers. Digitization is helping to manage limitations caused by institutional or financial obstacles that prevent access to historical biocollections (Hendrick, 2020, p. 1). This ongoing push to create digital surrogates allows for data to be gathered, shared, and analyzed much more efficiently and rapidly (Ridge, 2014, p. 4). NHCs form the historical baselines of a diverse range of species (Hendrick, 2020, p. 1). The online mobilization of their biodiversity data, into accessible digital content, is expanding the use of these collections across disciplines (2020, p. 1). Digitized NHCs are helping researchers to understand changes in species distribution and abundance brought on by anthropogenic forces (Miller-Rushing, 2020, p. 19). Interest in climate change, land-use, invasive species distribution, and ecological conservation have fostered this movement (2020, p. 19). In short, the digitization of NHCs is increasing the value of their physical counterparts exponentially (Hendrick, 2020, p. 1).

Data derived from digitized NHCs can illustrate changes taking place over time and across large areas (Mika, 2017, p. 60). For example, North American herbaria have collections of specimens that were gathered by botanists between 1870 to 1940 (Miller-Rushing, 2020, p. 19). By comparing the abundance and distribution of these specimens with the current populations, researchers are able to document the decline of endangered species and the spread of invasive species (2020, p.19). Similarly, the flowering dates of herbarium specimens compared with modern-day observations can detect shifts in bloom times brought on by climate change (2020, p. 19). It is through the digitization of NHCs that researchers are able to unlock vast amounts of biodiversity data and predict future outcomes.

Figure 2. The U.S. Herbarium's digitization conveyor belt. [Photo courtesy of National Museum of Natural History, 2021].

Crowdsourcing Natural History Collections

Natural history museums (NHMs) house vast collections of biodiversity specimens that were collected across centuries (Sforzi, 2018, p. 438). These collections include historical records which can aid in the understanding of biodiversity and anthropogenic changes taking place today (Hill, 2012, p. 220). The majority of NHM specimens are accompanied by hand-written or typed labels that detail information such as scientific name, location, date, and collector (2018, p. 438). Field notebooks and

correspondences may also be a part of a natural history collection (NHCs). These artifacts include raw data such as unpublished observations, species occurrence, habitat descriptions, climatological data, phenological records, illustrations, and travel logs (Mika, 2017, p. 59). Specimen labels, field notebooks, and correspondences contain a wealth of information and are considered to be as valuable as the biological specimens they represent (Alony, 2020, p. 18).

Figure 3. Assorted hand-written specimen labels. [Photo courtesy of Brian Stevenson, 2018].

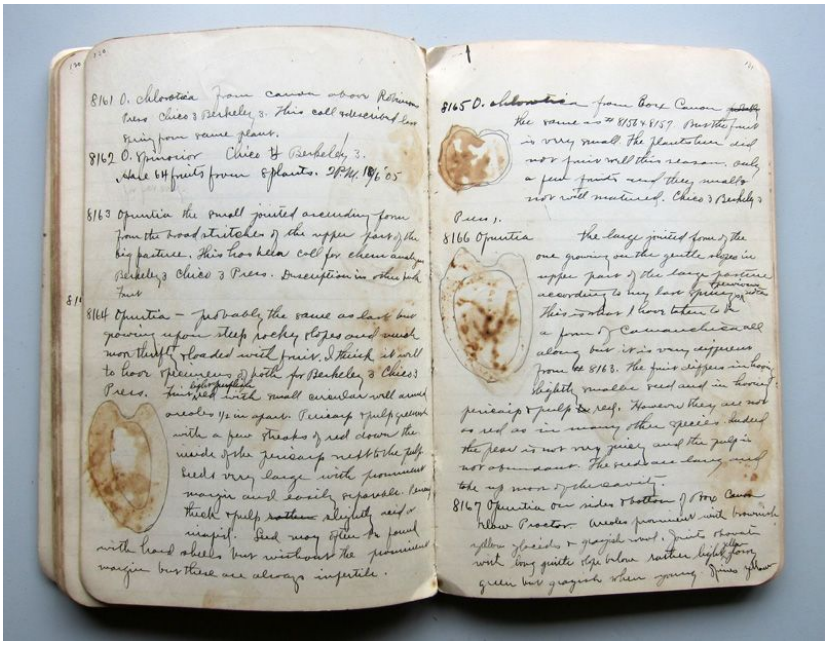

Figure 4. Pages from the Field Book Project. [Photo courtesy National Museum of Natural History, 2017].

The process of digitizing natural history collections (NHCs) includes multiple steps. The most time consuming, intensive and expensive part of the process is the transcription of biocollection-related text (Mika, 2017, p. 61). Online citizen-science crowdsourcing implores the aid of a massive network of amateur-experts to assist with transcription. Over time, projects both large and small have engaged millions of participants from around the world. Improving professional scientist's research and the means in which NHMs engage their audience. Online crowdsourcing is highly scalable, extending a citizen-science project's reach and engagement level (Sforzi, 2018, p. 439). By enlisting this method, NHMs are able to promote the value of working with citizen-scientists and deliver projects with strong scientific outcomes (2018, p. 444). This concept allows NHCs to be used in innovative ways that will promote them to a new generation of amateur-expert naturalists and ensure they will remain a source of research for years to come (Blaser, 2014, p. 46).

Online citizen-science crowdsourcing projects ask the public to undertake tasks that cannot be completed by computers. Handwritten historical records are extremely difficult, or impossible, to processes automatically (Blaser, 2014, p. 49). Rendering techniques such as optical character recognition (OCR) useless (Hill, 2012, p. 221). Humans have the innate ability to recognize patterns, make judgements, and process particular types of information (Shirk, 2020, p.9). They also have additional knowledge and experience to aid in the interpretation of biodiversity data (2020, p.9). In most cases, amateur-experts working on transcription are capable of outperforming any state-of-the-art algorithm (2020, p.9).

PLATFORMS

Zooniverse

Zooniverse was initially established in 2007 by a small group of academics for a single project, Galaxy Zoo (Mika, 2017, p. 68). The crowdsourcing platform encouraged amateur-experts to classify and describe the universe's galaxies, which had been previously been photographed by the Hubble Space Telescope (2017, p. 68). After its debut, Galaxy Zoo quickly gained the attention of a large network of citizen-scientists. Within the first 24 hours of it being announced, Galaxy Zoo received an average of 70,000 classifications per hour (Adams, 2012). Galaxy Zoo's popularity, as well as the amount of data being generated, was unexpected. The founding members of the platform were elated with what they had created- the "world's most powerful pattern-recognizing super computer" through the "linked intelligence" of engaged volunteers (2012).

From its humble beginnings, Zooniverse has grown to become the world's leading crowdsourcing citizen-science platform with over 2-million (2,333,567) registered volunteers that have completed in excess of 600-million (606,020,033) classifications (Zooniverse, 2021). Whereas Galaxy Zoo, part of the larger Zooniverse solar system, focused solely on astronomy; Notes from Nature hosts natural history projects ranging from ecology to biology. Notes from Nature invites amateur-experts to aid in the transcription of historical records such as specimen labels and field notebooks. These transcriptions help to further widen the pipeline of biodiversity data being delivered to researchers; spotlights participating institution's natural history collections (NHCs), and connects volunteers from around the world (Hill, 2012, p. 230). The scientific research conducted by Zooniverse has resulted in over 30 peer-reviewed publications and a handful of ground-breaking discoveries (Zooniverse, 2021).

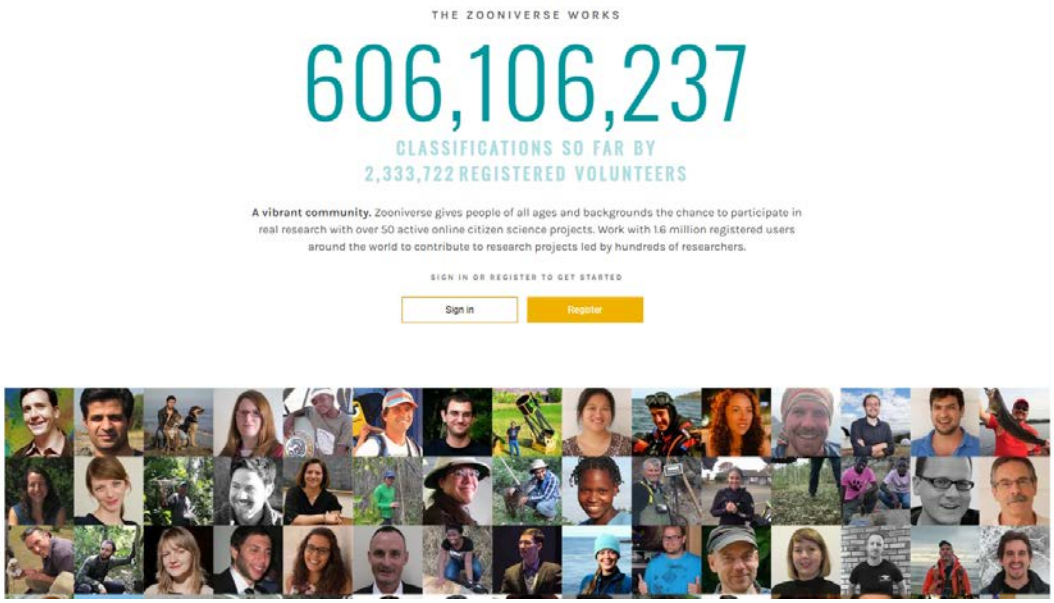

Figure 5. Zooniverse homepage. [Photo courtesy of Zooniverse, 2021].

Zooniverse's website is both organized and inviting. Citizen-scientists are directed along by tabs found at the top of the webpage, and illustrated tiles on the homepage identifying project genres.

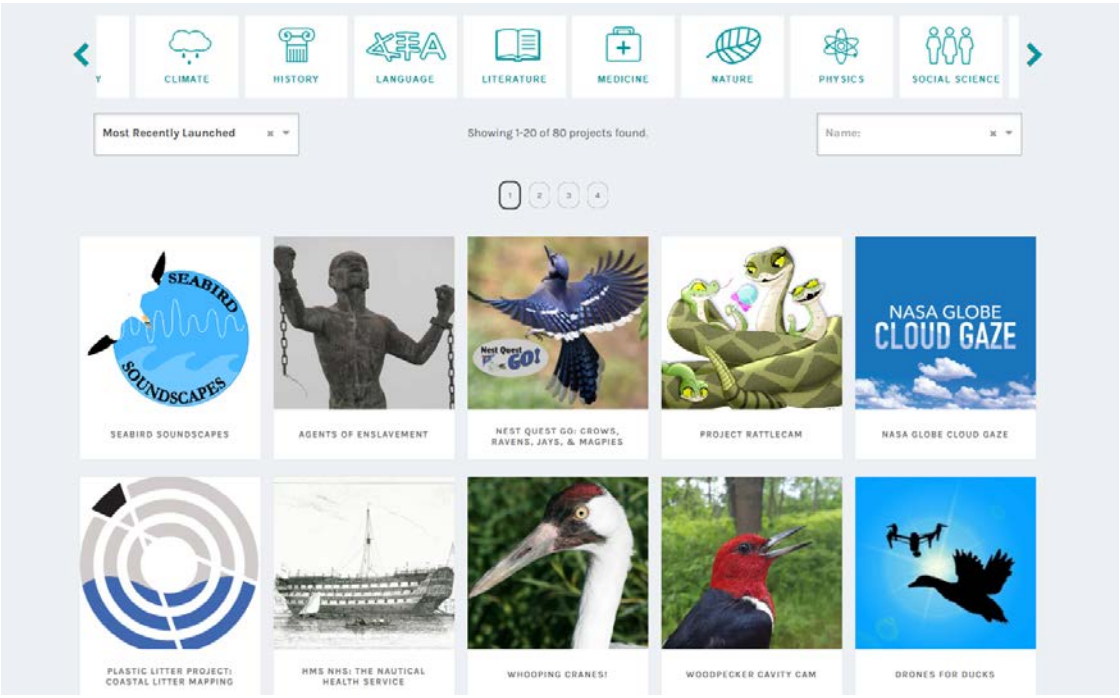

Figure 6. Zooniverse's project listings. [Photo courtesy of Zooniverse, 2021].

In order to transcribe a collection, citizen-scientists are encouraged to sign up for a free Zooniverse account by submitting their email address. Once they become a member, they're assigned a personal profile page which displays mission stats and favorite specimens (Hill, 2012, p. 225).

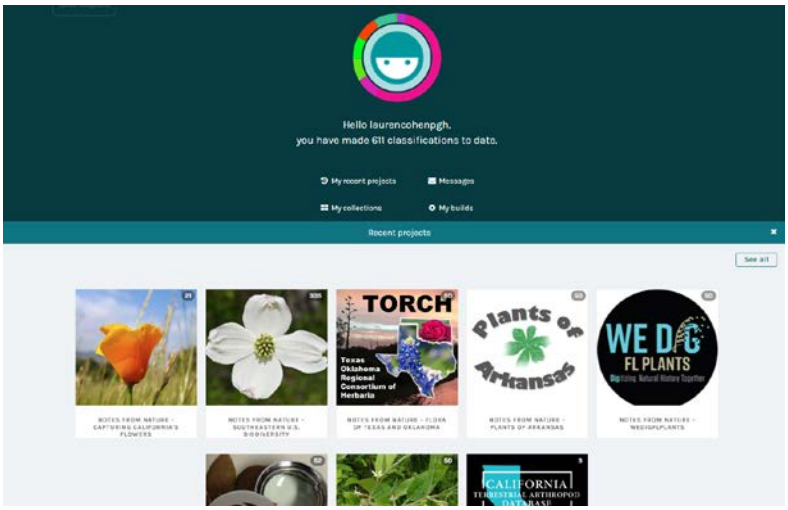

Figure 7. Zooniverse personal profile page. [Photo courtesy of Zooniverse, 2021].

Every project that is available on Zooniverse features a tutorial that appears before a citizen-scientist can begin transcribing. These tutorials are straightforward, illustrated, and easy to reference repeatedly throughout the transcription process; however, it is unfortunate that they are unable to be printed.

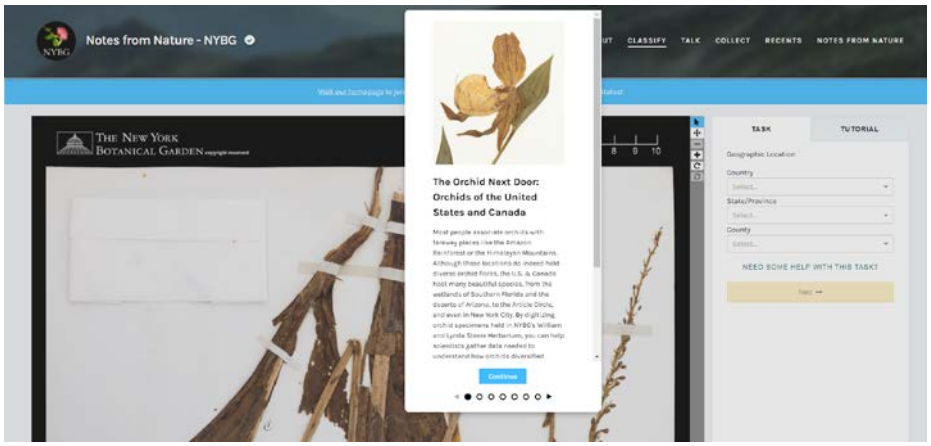

Figure 8. Tutorial for the NYBG Orchid Next-Door project. [Photo courtesy of Zooniverse, 2021].

Zooniverse relies heavily on "micro-tasking" to break up transcriptions that require a high degree of concentration (Mika, 2017, p. 68). By splitting tasks into more "manageable chunks with varying degrees of difficulty," user fatigue or boredom is eliminated and reliable data is produced (2017, p. 68). Zooniverse also recognizes and rewards the hard work of its citizen-scientists by presenting them with badges for certain milestones such as the completion of a certain number of transcriptions (Hill, 2012, p. 225). This reward systems enhances the user's interaction with a project and promotes future engagement (2012, p. 228).

A total of three replicate transcriptions are collected by Zooniverse for every transcribed record (2012, p. 226). This process, in addition to the validation of these records by professional scientists, ensures high quality data is being produced (Kosmala, 2016, p. 556). Zooniverse encourages its citizen-science community to interact with experts, or each other, via the Talk Board (Zooniverse Talk). The Talk Board allows volunteers to feel comfortable, seek help when needed, and make suggestions about how to improve the crowdsourcing platform. There are a massive number of conversations taking place daily on the platform. Not every available crowdsourcing platform facilitates this type of communication, which is why Zooniverse's citizen-science community has grown to this level (Kearney, 2012, p. 4).

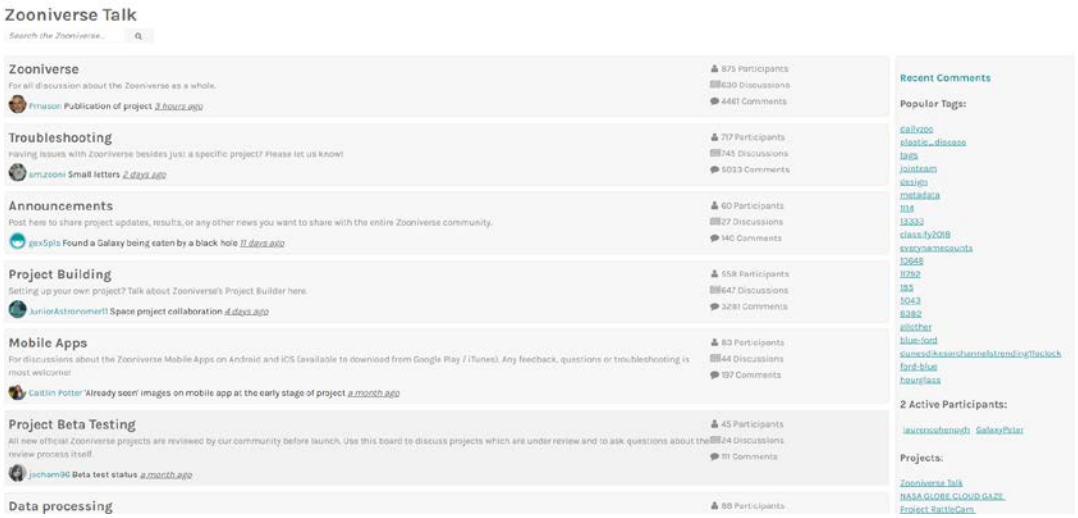

Figure 9. Zooniverse Talk. [Photo courtesy of Zooniverse, 2021].

DigiVol

The Atlas of Living Australia (ALA), in conjunction with the Australian Museum, launched DigiVol in 2011 (Alony, 2020, p. 17). This crowdsourcing platform is incredibly popular, and widely replicated by institutions from around the world including the Smithsonian Institution, Kew Gardens, and the New York Botanical Gardens (Mika, 2017, p. 70). DigiVol has won several international awards and processed an unprecedented amount of Australia's biodiversity data (2017, p. 70). DigiVol was built specifically for transcribing specimen labels, field notebooks, survey sheets, registers, and photographs (DigiVol, 2021). As of today, there are over nine thousand (9,842) active citizen-scientists transcribing on DigiVol who have completed upwards of 3-million (3,645,886) tasks (DigiVol, 2021).

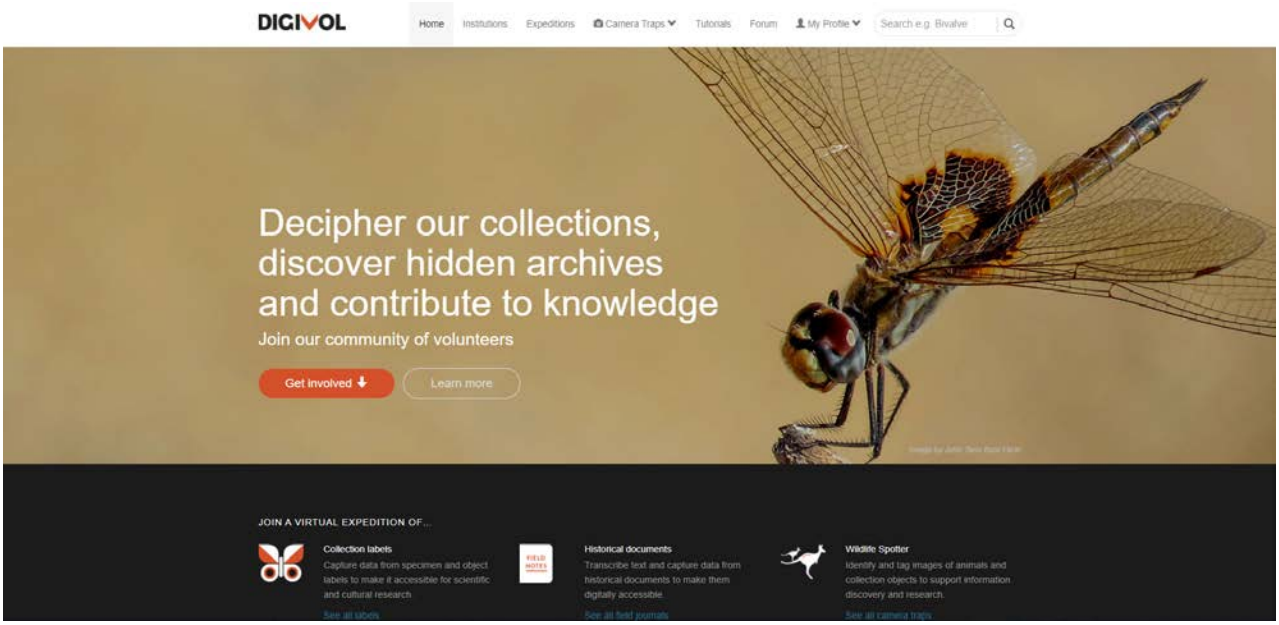

Figure 10. DigiVol homepage. [Photo courtesy of DigiVol, 2021].

Similarly, to Zooniverse, DigiVol is well organized and attractive; there are tabs along the top of the website and a search bar. Amateur-expert volunteers can navigate projects by clicking icons representing the main categories: Collection labels, historical documents, or wildlife spotter. There are

far fewer active projects on this platform than Zooniverse (known here as expeditions), yet there are more participating natural history museums (NHMs).

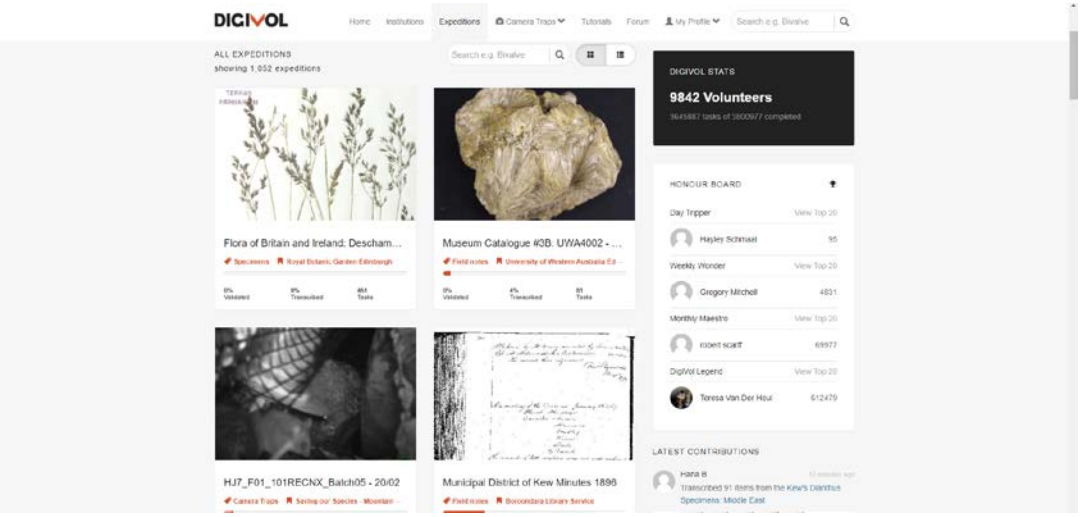

Figure 11. DigiVol's project listings. [Photo courtesy of DigiVol, 2021].

DigiVol also requires its users to enter an email address and register for a free account before transcribing a collection. A digital notebook tracks their personal progress across the platform.

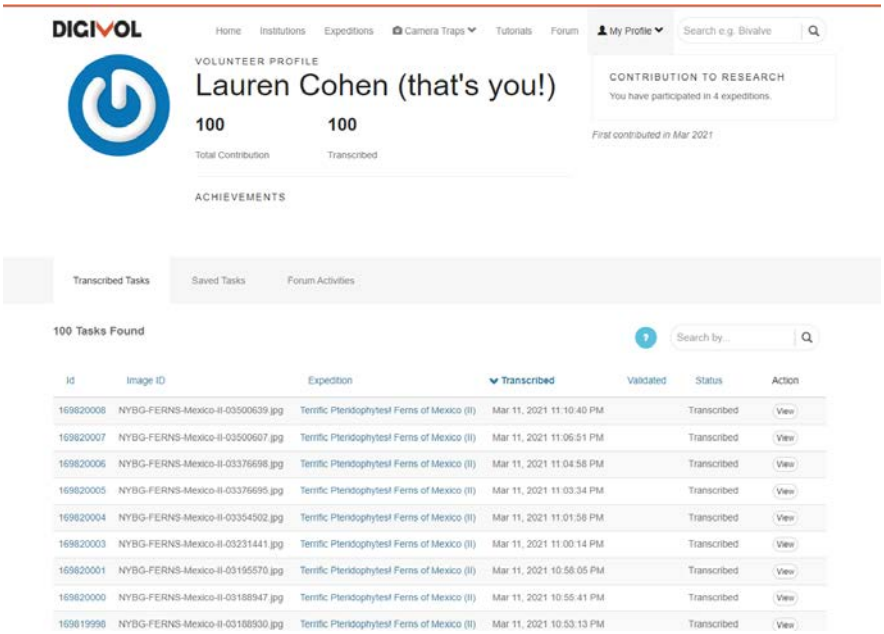

Figure 12. DigiVol personal notebook. [Photo courtesy of DigiVol, 2021].

DigiVol's tutorials are far superior to those found on Zooniverse. They are thoughtfully designed, project-specific, full of information, and incredibly helpful because DigiVol's transcriptions are quite intense. Digivol's transcription tasks are not broken into manageable bite-sized portions like Zooniverse, but require all the content to be entered at once. This takes quite some time to work through and can be very frustrating.

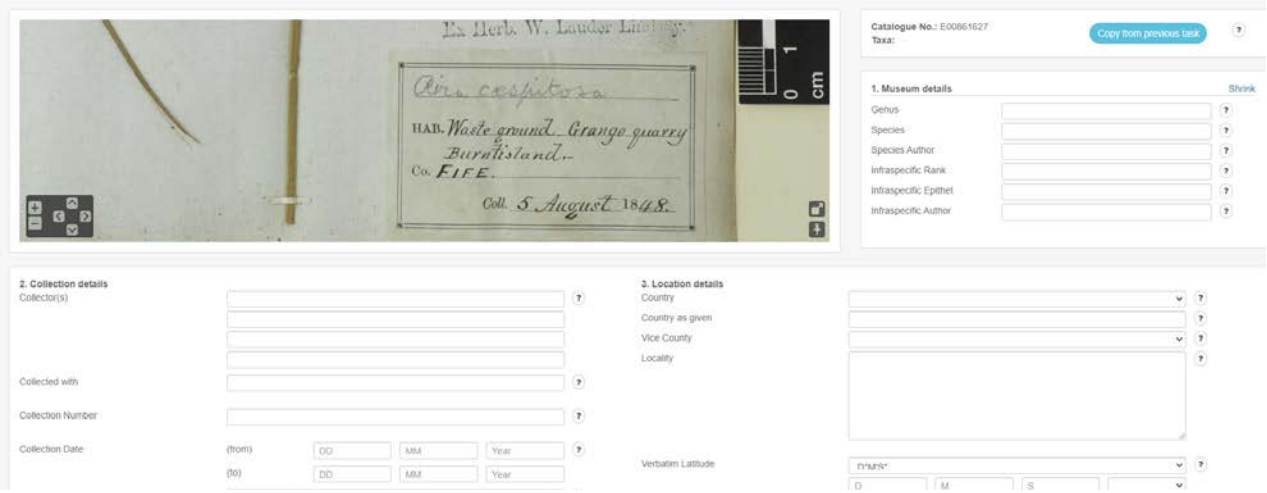

Figure 13. Sample of a DigiVol transcription. [Photo courtesy of DigiVol, 2021].

DigiVol's achievement awards and not provided as liberally as those on Zooniverse (Alony, 2020, p. 21). Top achievements are publicly announced along the right side of the website, and citizen-scientist's receive badges upon their ten-thousandth transcription (2020, p. 21). Once a transcription is submitted, the data is then checked by a validator. DigiVol's validators are assigned to the project by an administrator; and are usually an experienced volunteer who has shown, in the past, great attention to detail (Kearney, 2012, p. 4). Citizen-scientists are able to communicate with project administrators and fellow volunteers on a forum (Mika, 2017, p. 70).

Smithsonian Transcription Center

The Smithsonian Transcription Center was launched in 2012, and exclusively services the Smithsonian Institution's diverse complex of nineteen museums and nine research centers (Parilla, 2016, p.439). The crowdsourcing platform was developed as a way to increase access to the institution's digitized collections and modernize its engagement strategies (Smithsonian Transcription Center, 2021). With over 138 million objects and 153,000 cubic feet of archival materials, the Smithsonian has an abundance of data to share with researchers and the general public (2016, p. 439). Some of the natural history collections (NHCs) featured on the platform include specimen labels, photographs, correspondences, and field notebooks (Parilla, 2016, p. 439). Presently, there are over 50,000 actively engaged citizen-scientists (or "volunpeers") which have contributed to an ever-growing number (883,836) of transcriptions (Smithsonian Transcription Center, 2021).

The SI Transcription Center's website appears to be a little dated-looking compared to those of Zooniverse and DigiVol. There is also a lot going on at one time- many of the images move and there is an overwhelming amount of content (such as text and images) displayed on the homepage.

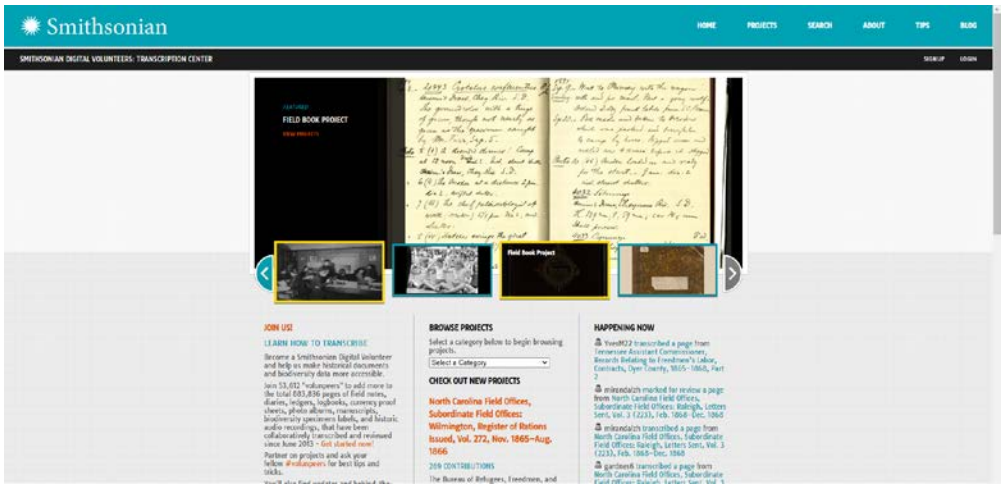

Figure 14. Smithsonian Transcription Center homepage. [Photo courtesy of Smithsonian Institution, 2021].

Despite the fact that the website appears to be out-of-date, the projects shared on the platform are by far the most exciting. They are also regularly promoted by the Smithsonian through a network of linked social media accounts and hashtags. Zooniverse and DigiVol do not partake is these outreach activities, yet their volunteer numbers never suffer as a result. Another exciting feature of the SI Transcription Center is a scale which rates each project's difficulty level from 1-5. One being for first-time transcribers, and five advanced projects.

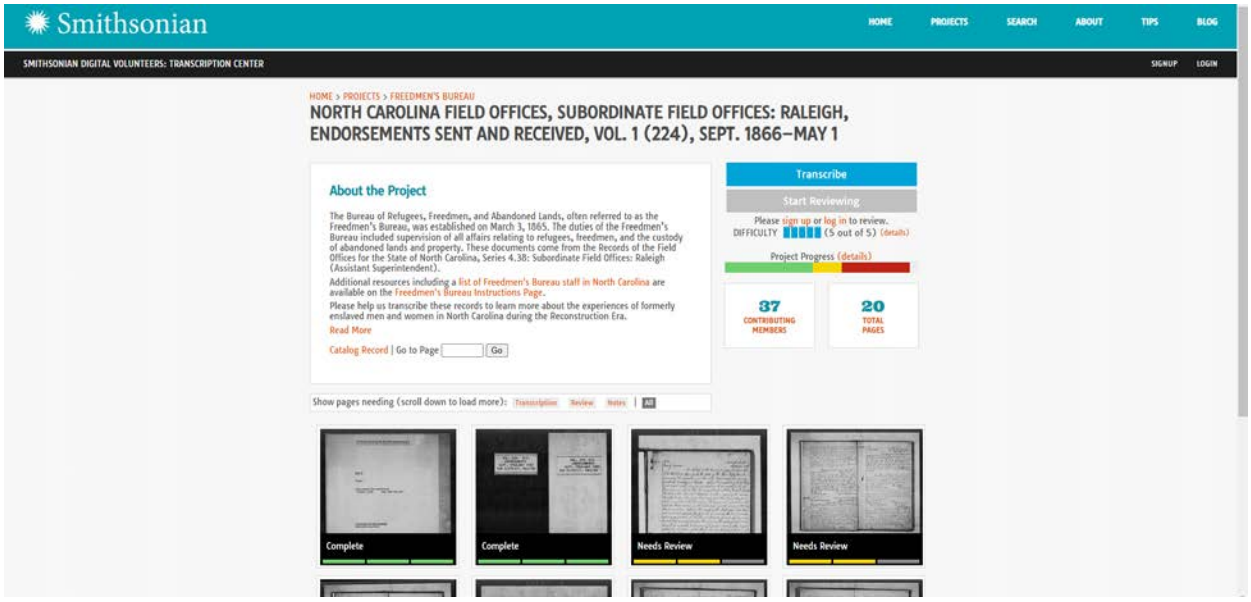

Figure 15. Smithsonian Transcription Center project layout. [Photo Courtesy of Smithsonian Institution, 2021].

The SI Transcription Center project's workflow is quite similar to that of Zooniverse, featuring small and manageable tasks. However, the way the projects are displayed on the website and corresponding tutorials are on par with that of DigiVol. When transcribing, citizen scientists can reference the digitized image on the left and data fields on the right; project-specific instructions remain accessible via a button located at the top of the page (Parilla, 2016, p. 444). Unlike both aforementioned platforms, users do not need to sign up for a free account in order to transcribe; there are also no awards given for participation and no forum to engage the community, just a blog that publishes information

about the active projects. To communicate with active amateur-experts, project coordinators or managers reach out via email.

The SI Transcription Center's validation process is also quite different than the other two platforms. The Smithsonian implores the use of a three-step peer-review process for all crowdsourced workflow (Parilla, 2016, p. 445). Any citizen-scientist can validate the transcribed pages of a project, just as long as they have an active account; but the last and final approval comes directly from Smithsonian staff (2016, p. 445)

CONCLUSION

Benefits: Biodiversity Data and Engagement

There are many contributing factors as to why natural history museums (NHMs) may want to digitize their biodiversity collections and share them online. The most relevant being: to make these natural history collections (NHCs) more accessible to researchers and connect with a wider audience. Citizen-science crowdsourcing can maximize the potential of both NHC accessibility and user engagement by generating vast amounts of biodiversity data in a more rewarding manner.

The digitization and crowdsourcing of NHCs are making biodiversity data more readily available to researchers. As NHMs often have extensive collections, online crowdsourcing is being more widely accepted as a suitable practice for improving access to historical records that may otherwise remain underutilized or unseen (Blaser, 2014, p. 48). Presently there are as many as 1-billion biocollections found in the United States locked away in cabinets, not yet digitized (Ellwood, 2015, p. 387). Resource-strapped museum professionals are incapable of processing the volume and breadth of this data singlehandedly (Mika, 2017, p. 61). With the mobilization of digital citizen-scientists, much needed aid is being delivered to participating institutions attempting to transcribe their NHCs. Digitized specimens, in association to crowdsourced historical records, represent vital benchmarks required to analyze the biological impacts brought on by anthropogenic change (Hill, 2012, p. 220). The research resulting from this type of online collaboration has been crucial in the study of invasive species, biological conservation, land management, pollination, phenological responses to climate change, species distribution and discovery (2012, p. 220).

Additionally, crowdsourced biodiversity research projects are leading to improved scientific literacy (Lepczyk, 2020, p. 1). Many of the participating amateur-experts engaged in transcription expand their understanding of NHCs and develop a deeper appreciation for science in general (2020, pp. 1-2).

This educational component of crowdsourcing offers NHMs an opportunity to effectively disseminate their core values. NHM missions often focus on the development and accessibility of their collections; as well as the advancement of scientific knowledge, public engagement, and safeguarding of the natural world (Sforzi, 2018, p. 441). Through the use of citizen-science crowdsourcing projects, institutions are able to meet their mission-related objectives and expand their reach. This symbiotic relationship is mutually beneficial because it strengthens collections and facilitates discovery (Mika, 2017, p. 61). C

Crowdsourcing provides an extraordinary opportunity for users to gain first-hand experience with NHCs in meaningful ways (Ellwood, 2015, p. 384). Instead of a one-dimensional experience of browsing collections online, citizen-science projects invite the audience to actively participate in the research process (Owens, 2014, p. 279). This opportunity allows for the user to do more than consume information; emphasizing the fundamental reason why NHCs are being shared online, to improve access and engage with non-traditional audiences (2014, p. 279).

Concerns: Data Quality

Despite its many benefits, crowdsourcing still incites skepticism across the professional science community (Kosmala, 2016, p. 551). This negativity and distrust stem from the fact that citizen-science is not considered to be a mainstream approach to research, and is therefore questionable science (2016, p. 551). Additionally, many professionals believe that amateur-expert volunteers are not committed or skilled enough to perform at the same level as paid staff; resulting in the doubting of their abilities and motivations to produce reliable data (2016, p. 551). It is important to emphasize that citizen-science is not a replacement for professional scientists (Lepczyk, 2020, p. 1). There will forever remain a demand for highly trained individuals who have pursued higher-education and have professional experience in scientific research (2020, p. 1). However, it is worth recognizing the valuable contributions amateur-expert naturalists have made over the last few centuries.

With a growing number of peer-reviewed publications, citizen-science is progressively finding its way into acceptance by proving that these projects can produce data that is equivalently accurate to that of the professionals (Kosmala, 2016, p. 551). While citizen-science projects can vary in subject matter, task, and scale, the production of reliable data remains a constant focus (2016, p. 551). Effective methods for acquiring quality results vary based on the type of data being produced and the resources available to the citizen-science project (2016, p. 553). Most of these procedures are identical to those used by the professional scientific community; therefore, crowdsourced projects should be judged on their methods and stewardship versus the fact that amateur-experts were part of the research process (2016, p. 552).

The most successful crowdsourcing projects rely on a suite of methods to boost data accuracy including: iterative project development, volunteer training, and expert validation (Kosmala, 2016, p. 551). Developing tools and protocols for a citizen-science projects require iteration; and the utilization of pilot or beta testing ensures tasks will be performed successfully without confusion or error (2016, p. 553). The individual accuracy of a citizen-scientist varies depending on task difficulty and experience level (2016, p. 553). In most cases, accuracy improves with proper instruction, simpler tasks, and further project participation (2016, pp. 553-554). Lastly, to ensure high-quality data is produced expert validation is required. This verification process can either be executed by professionals or seasoned citizen-scientist volunteers (2016, p. 556).

Discussion

The following observations have emerged from the literature review and case studies accompanying this research project. Moving forward into the 21st century, natural history museums (NHMs) will require new methods for unlocking biodiversity data and engaging their audiences. Citizen-science crowdsourcing is a sustainable model for NHMs eager to make their collections more accessible online. Crowdsourcing is a tool that can cost effectively collect and analyze historical records associated

with digitized specimens; therefore, enabling biodiversity data to be gathered and shared much more rapidly across the research community. Despite this, online citizen-science projects hosted by natural history institutions are still in their infancy and our knowledge about their full potential is still not entirely understood.

Citizen-science crowdsourcing projects continue to grow online in size, scope, and significance. Platforms such as Zooniverse, DigiVol, and the Smithsonian Transcription Center are the most popular on the web because of their proven track record for producing high-quality data, diverse projects, and large networks of volunteers. These platforms are successful because they do more than invite the users to donate their time and expertise. In their own unique ways, each platform cultivates an environment that encourages participation between amateur-experts and professionals that is mutually beneficial. Simply generating and communicating scientific knowledge is no longer a sufficient way for NHMs to interact with their audiences. These platforms are helping to usher in the new dawn of citizen-science by breaking down barriers between the public and professional scientific community; and revitalizing the age-old tradition of cross-disciplinary collaboration that past scholars once encouraged.

The benefits associated with citizen-science crowdsourcing extends far beyond that of data collection or analysis. These platforms act as a tool for education and public engagement; and hold great promise to enhance scientific literacy due to their participatory hands-on nature. For the skeptics, rather than becoming too concerned and reactive with data quality, focus should be directed towards improving project design which can serve as a stabilizing force to enhance accuracy. There are significant number of citizen-science generated scholarly publications produced via these platforms to support that claim.

Further research needs to be conducted to better understand the reasonings behind the lack of crowdsourcing use by NHMs. This researcher was unable to reach any conclusions-

REFERENCES

Adams, Tim. (2012). Galaxy Zoo and the New Dawn of Citizen Science. The Guardian. Retrieved from https://www.theguardian.com/science/2012/mar/18/galaxy-zoo-crowds Owens, Trevor. (2014). Chapter 12: Making Crowdsourcing Compatible with the Missions and Values of Cultural Heritage Organizations. In Crowdsourcing our Cultural Heritage. Ashgate Publishing: New York, New York.

Alony, I., Haski-Leventhal, D., Lockstone-Binney, L., Holmes, K. & Meijis, L. C. P. M. (2020). Online volunteering at Digivol: an innovative crowd-sourcing approach for heritage tourism artifacts preservation. Journal of Heritage Tourism, 15 (1), 14-26. Retrieved from https://www.tandfonline.com/doi/abs/10.1080/1743873X.2018.1557665

Blaser, Lucinda. (2014). Chapter 2: Old Weather, Approaching Collections from a Different Angle. In Crowdsourcing our Cultural Heritage. Ashgate Publishing: New York, New York.

Burnett, Linda Anderson. (2021). Eighteenth Century Citizen Science Inspired by Linnaeus. Knut and Alice Wallenberg Foundation. Retrieved from https://kaw.wallenberg.org/en/research/eighteenth-century-citizen-science-inspired-linnaeus

DigiVol. (2021). About DigiVol. DigiVol. Retrieved from https://volunteer.ala.org.au/about/index#/what-is-digivol

Ellwood, E., Dunckel, B., Flemons, P., Guralinik, R., Nelson, G., Newman, G., & Mast, A. (2015). Accelerating the Digitization of Biodiversity Research Specimens through Online Public Participation. *BioScience, 65*(4), 383-396. Retrieved from https://www.jstor.org/stable/90007259

Hecker, S., Haklay, M., Bowser, A., Makuch, Z., Vogel, J., & Bonn, A. (2018). Citizen Science: Innovation in Open Science, Society and Policy. UCL Press: London, England.

Hedrick, B., Heberling, M., Meineke, E., Turner, K., Grassa, C., Park, D., Kennedy, J., Clarke, J., Cook, J., Blackburn, D., Edwards, S. & Davis, C. (2020). Digitization and the Future of Natural History Collections. BioScience. 10.1093/biosci/biz163.

Hill, A., Guralnick, R., Smith, A., Sallans, A. Gillespie, R., Denslow, M., Gross, J., Murrell, Z., Conyers, T., Oboyski, P., Ball, J., Thomer, A., Prys-Jones, R., de la Torre, J., Kociolek, P. & Fortson, L. (2012). The notes from nature tool for unlocking biodiversity records from museum records through citizen science. ZooKeys 209, 219-233. Retrieved from https://www.researchgate.net/publication/230616771_The_notes_from_nature_tool_for_unlocking_biodiversity_records_from_museum_records_through_citizen_science

Kearney, Nicole. (2015). Transcribing Between the Lines: Crowdsourcing Historic Data Collection. Museums and the Web, Asia. Retrieved from https://mwa2015.museumsandtheweb.com/paper/transcribing-between-the-lines-crowd-sourcing-historic-data-collection/

Kosmala, M., Wiggins, A., Swanson, A. & Simmons, B. (2016). Assessing data quality in citizen science. Frontiers in Ecology and the Environment, 14(10), 551-560. Retrieved from http://www.jstor.org/stable/24891216

Lepczyk, C. A., Boyle, O. D., & Vargo, T. L. V. (2020). Introduction. In Handbook of Citizen Science in Ecology and Conservation. University of California Press: Oakland, California.

Mahr, D., Gobel, C., Irwin, A., & Vohland, K. Chapter 7: Watching or Being Watched. (2018). In Citizen Science: Innovation in Open Science, Society and Policy. UCL Press: London, England.

Mika, K., DeVeer, J. & Rinaldo, C. (2017). Crowdsourcing Natural History Archives: Tools for Extracting Transcriptions and Data. Biodiversity Informatics, 12, 58-75. Retrieved from https://core.ac.uk/reader/235873782

Miller-Rushing, A. J., Primack, R. B., Bonney, R., & Albee, E. (2020). Chapter Two: The History of Citizen Science in Ecology and Conservation. In Handbook of Citizen Science in Ecology and Conservation. University of California Press: Oakland, California.

Parilla, L. & Ferriter, M. (2016). Social Media and Crowdsourced Transcription of Historical Materials at the Smithsonian Institution: Methods for Strengthening Community Engagement and Its Tie to Transcription Output. The American Archivist, 79 (2), 438-460. Retrieved from http://www.jstor.org/stable/26356669

Ridge, Mia. (2014). Introduction. In Crowdsourcing our Cultural Heritage. Ashgate Publishing: New York, New York.

Sforzi, A., Tweddle, J., Vogel, J., Lois, G., Wagele, W., Lakeman-Fraser, P., Makuch, Z., & Vohland, K. (2018). Chapter 29: Citizen Science and the Role of Natural History Museums. In Citizen Science: Innovation in Open Science, Society and Policy. UCL Press: London, England.

Shirk, J. L. & Bonney, R. (2020). Chapter One: What is Citizen Science? In Handbook of Citizen Science in Ecology and Conservation. University of California Press: Oakland, California.

Smithsonian Transcription Center. (2021). About: What is the Transcription Center? Smithsonian Institution. Retrieved from https://transcription.si.edu/about

Zimmer, Carl. (2011). Darwin Meets the Citizen Scientists. Discover Magazine. Retrieved from https://www.discovermagazine.com/planet-earth/darwin-meets-the-citizen-scientists

Zooniverse. (2021). Homepage. Retrieved from https://www.zooniverse.org/