

Opening up the Archive: An Approach to Volunteer-led Crowdsourcing in the Museum and Online

Authors

Geoffrey Belknap¹, Samantha Blickhan², Alex Fitzpatrick³, Lynn Wray⁴

Abstract

This paper outlines a new, collaborative approach to working with volunteers from local communities to bridge the gap between in-person and online volunteering, and ultimately open up photographic archives. Based on a case study at National Science and Media Museum in Bradford, UK, we outline the process by which local volunteers have created a Zooniverse project using images that they have selected and digitised from a large mid-20th century photographic archive related to African Caribbean history. The paper will summarise the development of this project and outline the process we have developed from volunteer-led digitisation to crowdsourcing, with the goal of creating a set of tools and processes for other cultural heritage organisations to democratise access to their photographic collections.

Keywords

Digital Humanities; Citizen Science; Heritage Volunteer; Participatory Practice; Digitally Enabled Participation

Introduction

In cultural heritage organisations across the UK and US, volunteers are a critical part of institutional infrastructure. Volunteer labour is ubiquitous, spanning disciplines from art museums to planetaria. Within cultural organisations, volunteers interact with visitors, work with collections, and provide administrative support, among myriad other tasks that vary across institutions. Volunteers help to ensure that organisations run smoothly, even when budget limitations prevent the hiring of additional staff. At the same time, volunteers personally benefit from their association with museums, either due to interest in the content of collections, a feeling of 'giving back' to one's community, opportunities for socialisation, or career-oriented learning and experience-building.

This paper outlines the approach developed by the Arts and Humanities Research Council funded project, *Communities and Crowds: Expanding Volunteer Programs*

¹ Keeper of Science and Technology, National Museums Scotland

² Zooniverse Co-Director & Humanities Lead, Adler Planetarium

³ Research Associate, National Science and Media Museum

⁴ Research Fellow, University of Leeds

Across Physical and Digital Spaces for Cultural Institutions, to create a volunteer-led approach to crowdsourcing, with additional support from new custom tools and infrastructure on the Zooniverse crowdsourcing platform. The approach detailed in this article outlines the midway findings of the project, and details how cultural heritage institutions can create space for volunteers to decide what gets digitised, how it is digitised, and how the digital assets are utilised to create a volunteer-led crowdsourcing platform.

Review of literature

Recent scholarship in digitally-enabled participation has been questioning the efficacy of online crowdsourcing projects which do not engage with the inherent power structures present in efforts led by institutional researchers (Jancovich and Stevenson 2023; Mahr and Dickel 2019). As cultural institutions increasingly make use of digital technology to engage audiences, the need for volunteers has arisen in those spaces as well. However, digital volunteer opportunities often focus on volunteers as audiences, rather than as an essential part of the institutional team. One example of attracting digital volunteers to heritage institutions is through online crowdsourcing projects. Crowdsourcing is a form of digitally-enabled participation which invites volunteers to engage with institutional collections by performing meaningful tasks (Ridge et al. 2021). It is often viewed as a way to simultaneously encourage public engagement with heritage materials, while completing essential institutional data processing tasks that can result in making collections searchable and discoverable by a wider range of audiences. However, most crowdsourcing projects are created 'offline' and only shared with online volunteers once the project's aims and research questions are fully formed. In this way, digital volunteers are assigned the role of 'users' or 'audience', rather than as co-creators or collaborators.

In this effort, we aim to re-examine the role of the volunteer in the physical and digital realms, with the ultimate goal of breaking down hierarchical barriers and providing a roadmap for teaching cultural institutions to use online engagement as a model for interacting with both digital and in-person volunteers as collaborators with institutional staff, rather than as 'the ultimate frequent visitors' (Orr 2006). In particular, we focus on how this practice can be realised through the process of online crowdsourcing, using the Zooniverse platform as a case study.

Communities and Crowds brings about a new model for understanding how to engage with museum audiences through a connected experience of in-person and online volunteer engagement. Projects working with volunteer communities in person and online are typically developed and evaluated as discrete activities. Models of participatory engagement with museums (Benoit III and Everleigh, 2019; Simon, 2010) focus on the pathways for deep in-person engagement, while forms of online engagement such as crowdsourcing (Ridge 2014; Tauginienė et al. 2020) focus on discrete forms of digital engagement with heritage collections.

Through this experimental project, we explore and document how levels of volunteer engagement and agency with heritage collections can be amplified by incorporating

volunteer communities throughout the entire process of developing and facilitating a crowdsourcing project—from image selection and digitization to project development, launch and moderation. Institutional volunteers and audiences are not mutually exclusive, which makes volunteer involvement in cultural institutions even more powerful when harnessed to its full potential. Current efforts in the UK and US to decolonize the archive speak to the need to include communities in the process of preserving, recording, and remembering their own culture and history (Brook and Neath 2018; Lydon 2017; Proctor 2014; Van Broekhoven 2018). We believe that this process is strengthened by empowering volunteers to participate in institutional efforts from the start—to make decisions about what is digitised, how collections are presented, why certain items are prioritised—as collaborators and co-creators.

Methodology

A new model of volunteer engagement

Recruitment

The Science Museum Group (SMG) already has a large and diverse volunteering programme, and has ambition to be the leading national institution in the UK for volunteering by 2030 (Science Museum Group 2020). The National Science and Media Museum (as part of SMG) also has a desire to increase opportunities for participatory engagement with its local audiences. Drawing on learning from the recent AHRC funded research project, *Bradford's National Museum* (Bradford's National Museum Project 2023) which created the co-curated exhibition *Above the Noise: 15 Stories from Bradford* (Above the Noise 2019), it was clear that the volunteer selection for the *Communities and Crowds* project would need to involve local community groups that were representative of the cultural diversity of Bradford and would be keen to explore research interests related to heritage. Due to social distancing regulation at the time of development that greatly reduced the capacity of research space in the museum, recruitment was limited to a total of four volunteers.

We sent invitations to volunteers via local groups with interests in cultural heritage, with a focus on communities that were not engaged with the museum at the time. The volunteer roles were positioned as opportunities for local community members to become collections researchers and “mine” the archive to collaboratively develop a digital crowdsourcing project based on their own interests and heritage.

A majority of the resulting volunteer team would come from groups representative of the local African-Caribbean community, with a shared interest in amplifying stories of Caribbean heritage and histories of migration, as well as challenging racist attitudes and assumptions. We recruited four volunteers in total. Three volunteers, Sandra Rowe, Maureen Rowe and Lincoln Anderson were active members of the Bradford African and Caribbean community; and the fourth volunteer, Rebecca Smith, was a previous employee of NSMM and PhD student, who joined the project due to her long experience and passion of working with photographic collections. The volunteers also shared an understanding that barriers existed that prevented

engagement between the museum and historically underrepresented and underserved communities; as such, there was a shared notion amongst the volunteers of this project as a way of creating spaces of dialogue and engagement that were open and accessible for their communities.

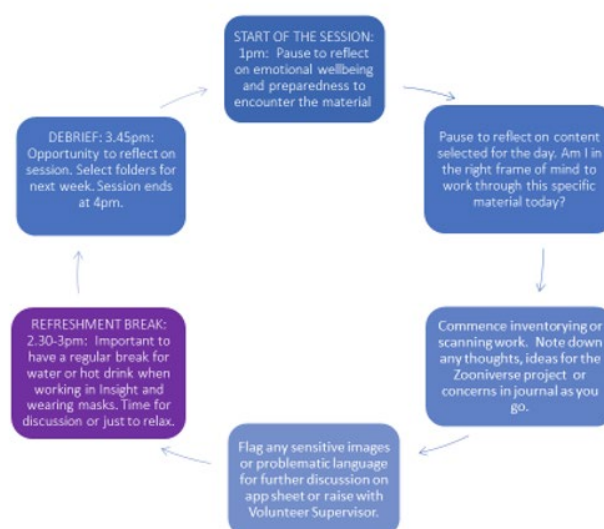
Staff Training

Throughout the process of establishing the volunteer team, we were aware that as the employed project team members, we needed to have concrete tools to address the potential impact of working with photographic collections that contained problematic imagery, content and language for our volunteers and volunteer supervisors. In order to support our project staff in creating a safe space for everyone involved in the project, and a better understanding of how to mitigate any potentially traumatic experiences for our volunteers, we undertook training from the Black, African and Asian Therapy Network (BAATN) (BAATN 2023). The training with BAATN included addressing how we create a safe and supportive environment for staff and volunteers to bring up any negative experiences or feelings as they were working through the documentation and digitization of the collections. One of the key outcomes from this training was a recommendation to pay particular attention during the volunteering sessions for time out, breaks, and collective reflection.

Session Planning Circuit

Given the particular research interests of the volunteer group around images representing Afro-Caribbean cultures, countries and experiences in the mid-20th century, we worked with BAATN to develop an approach for encountering sensitive content and problematic language in the archives. The outcome of this work was the creation of a session protocol, which we embedded into the volunteer workflow.

Session Plan – Volunteer Researchers (Wellbeing)



Session planning cycle for each volunteer session

This included the provision of introductory sessions to give volunteers further context of the archives they would be working with, allowing them to make an informed choice regarding their participation in the project. We also held discussions with the volunteers regarding physical and emotional responses to encountering challenging and potentially traumatic archival material, with possibilities raised for providing therapeutic support if needed (e.g., holding reflection sessions outside of the museum, signposting to BAATN resources and therapists).

Within the volunteer workflow itself, we made space for reflection and processing of subject matter. This consisted of time at the start of each session to ensure that each volunteer was in the right headspace to work on the archival material for the day, as well as regular break times for refreshments outside of the collections area. Each session would also end with a debrief and reflection in the museum cafe, providing the volunteers with the opportunity to bring up any concerns or discuss any issues that may need to be resolved for the next session.

Volunteer-Led Documentation and Digitization Process

One of the core principles of the project was creating a documentation and digitisation workflow which would follow best practice principles within a heritage setting, while also streamlining the process so that a volunteer without knowledge or expertise in these processes could develop confidence in creating digital assets without the demands or expectation of professional outputs. We were keen that the images and metadata that the volunteers created as part of this project would be able to be useful when they came to developing their Zooniverse project; while also creating standard data which could be incorporated into SMGs collections management system (MIMSY) and online collections platform.

To support this work, we developed augmented workflows for three discrete pieces of work that would be led by the volunteers for photographic images: storage and re-housing; barcoding and core documentation; and digitisation and post-processing. In order to create consistent workflows for these three processes, we collaborated with SMG staff from conservation, documentation and digitization teams respectively to adapt in-house processes to the volunteer workflow.

Future outputs for the *Communities and Crowds* project will share the specific tools developed for these new workflows via a Tool Shed model (see Conclusion and Next Steps below). However, the following is a brief description of the adapted workflow for each of the three processes.

Storage and Re-Housing.

The volunteers were working with a photographic archive that has a unique organisational structure and storage. The Daily Herald Archive is a ~3.5 million photographic archive from one of the largest 20th century national UK newspapers

(Daily Herald Archive 2022). The archive served as the image library for the newspaper and is currently stored in the same filing cabinets and filing structure created by the photographic editors. In order to retain the integrity of the organisational structure of the archive, the volunteers selected entire folders from the archive (rather than selecting images across the whole archive based purely on content). The first step after selecting the folder included taking the images out of their folder; placing them in individual melinex sleeves; barcoding each sleeve; and placing the sleeved photographs in a barcoded (at container level) acid-free box. The original folders (now empty of photographs) were then returned to their original filing cabinets.

Barcoding and core documentation.

The next step after re-housing the photographs was to create a core metadata entry for the images before scanning. The collections management system (CMS) used by SMG is called MIMSY, and requires controlled password and network access to update any entry. As direct access to the CMS by the volunteers would have created significant barriers for the volunteers and risks to organisation, the SMG data team created a flexible data entry interface for the volunteers to use based on a google tool called App Sheet (App Sheet 2023). The app sheet tool allowed us to replicate the filing structure of the archive and the data entry requirement for each photographic object. Working with this tool, the volunteers would first scan the barcode for the container that they were working on; then scan each photograph to create a new record; and finally a drop down menu asked them to enter core metadata for each photograph, which included adding a part number; identifying any hazards; identifying condition; identifying type of photograph; identifying intellectual property; identifying any named photographer; and three yes/no categories for oversized, potentially sensitive imagery, and/or problematic language.

Digitisation and post-processing.

Working in pairs, the volunteers would then digitise the prepared photographs. Using an Epson perfection V850 pro scanner, one volunteer cleaned the flatbed glass with an air blower and microfibre cloth, placed the image in the centre of the glass and a b/w QP colour correction card spaced below the image. Once the scanner was closed, the other volunteer would scan the image from the attached laptop into a prepared network storage file using the software, Adobe Bridge. The teams would replicate the process after flipping the photograph and scan the verso of the image.

Following the scanning, the volunteer at the laptop would use the barcode scanner to add the metadata to the digital asset. They would then crop the image and use the colour correction tool to balance the black and white contrast of the image. They would then save final versions of the verso and recto of each photograph in the networked drive.

Together, these three processes have allowed for the project team and volunteers to have consistent and reliable metadata and digital assets, which have been critical to the second stage of the project—creating the Zooniverse project.

Forward Planning for Zooniverse (Workshops)

Alongside the digitisation sessions, we also held workshops and discussion sessions with the volunteers to support the Zooniverse project planning process. These sessions were vital in ensuring that the interests and aims of the volunteers continued to remain at the forefront of the Zooniverse project development, despite the volunteers themselves not being directly part of the technical aspect of this part of the project. It should be noted that this was a decision made by the volunteers themselves; prior to the start of this process, they were given the opportunity to be more directly involved with the project build. However, the volunteers were more interested in creating an overall framework for the Zooniverse team to follow in developing the project.

Although this decision allowed the volunteers to contribute at their own comfort levels and led by their own interests, the downside was that the volunteers had some difficulty in visualising how the platform would work on a practical level. To help with this issue, a “paper prototyping” session was held, which simulated the experience of a Zooniverse project volunteer using previously selected photographs from the archive. Volunteers were asked to approach these images from the perspective of someone seeing them for the first time and attempt some of the proposed workflow tasks for Zooniverse, such as transcribing text and classifying photographs by subject and location. Afterwards, the volunteers discussed what sort of new information could be gained from these actions and how this ultimately aligned with their previously determined goals for the project.

Zooniverse Design/Infrastructure Development

While the NSMM volunteers worked on digitisation and project design, the Adler and Oxford Zooniverse teams carried out a series of custom development efforts. Our technical approach stemmed from the current assumptions and limitations of the Zooniverse platform, particularly surrounding spaces for volunteer interaction and the relationship between these spaces and the act of classifying. We also wanted to consider how creative technical changes could challenge the inherent power structures held up by the current technical infrastructure, e.g. placing the classification task (data production) above other modes of participation such as posting on project message boards (public engagement). We addressed these goals through a two-phase technical development effort.

1. QuickTalk

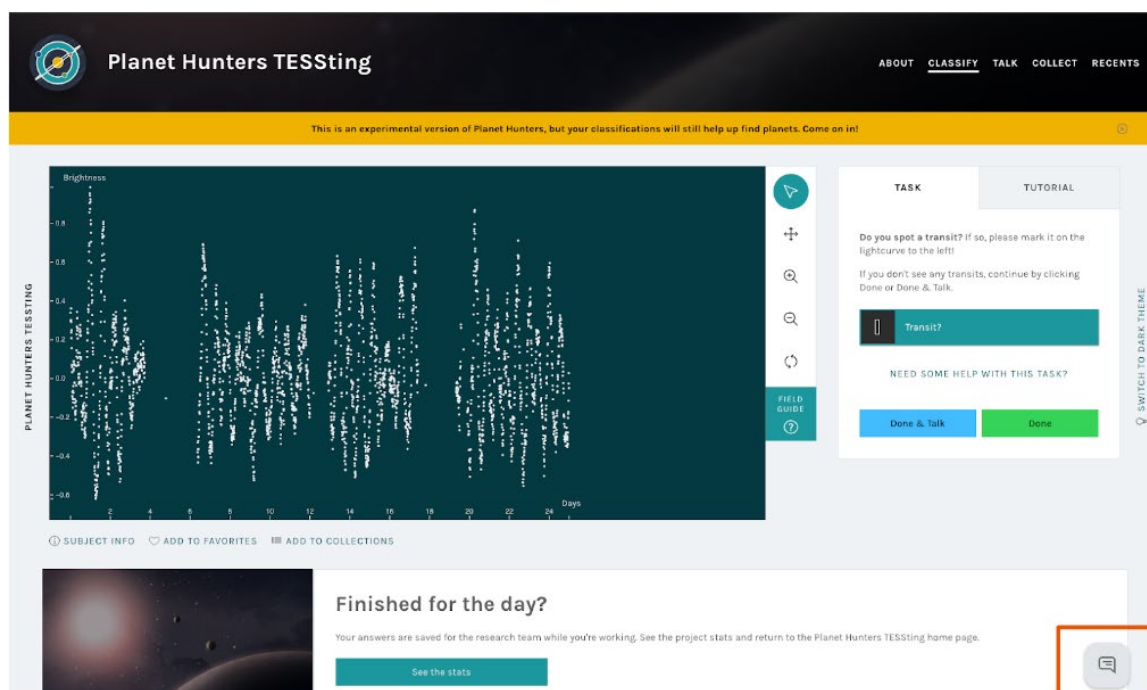
All Zooniverse projects come with their own message board, known as ‘Talk’, where participants can interact with one another as well as members of the project research team. Topics range from image-specific comments to troubleshooting discussions or

general content questions. The most common mode of accessing Talk is via the classification interface. After finishing a classification, volunteers have the choice to click either a button labelled 'Done' or one labelled 'Done & Talk'. 'Done' submits the classification and loads the next subject. 'Done & Talk' submits the classification and takes the participant directly to the message board page for that subject. Once there, they can read or post comments, ask questions, share discoveries, etc.

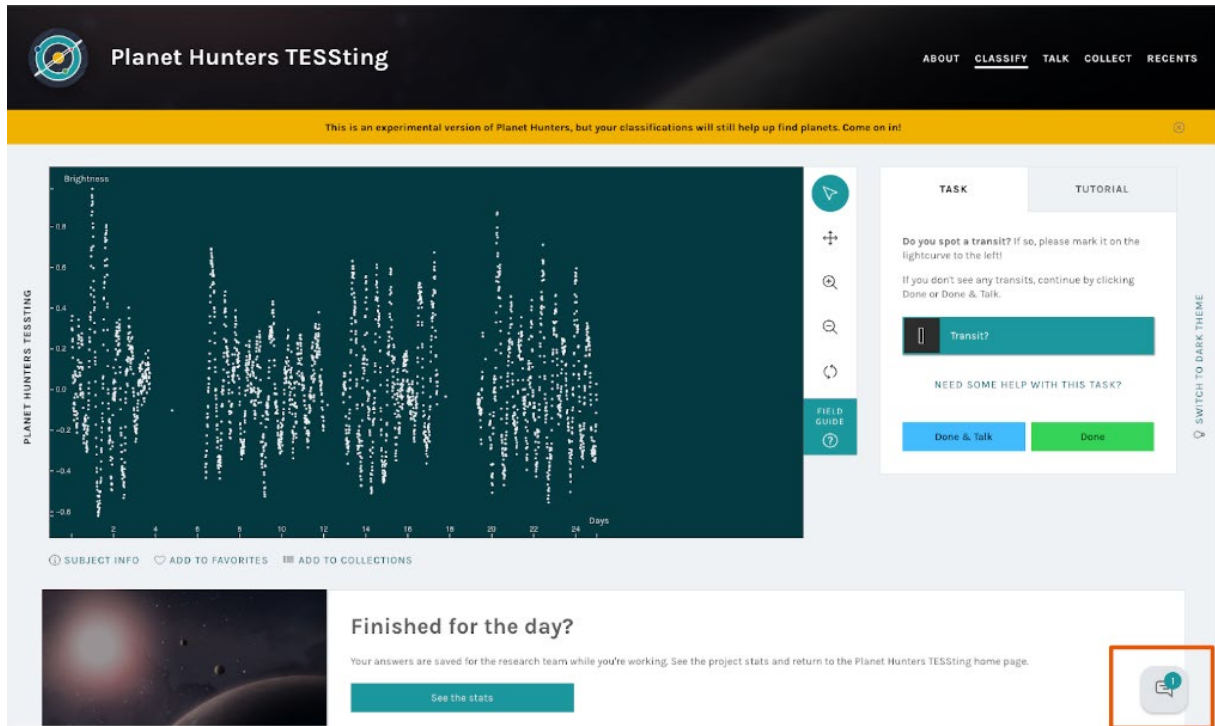


The Done & Talk and Done buttons, as shown on the Zooniverse classification interface.

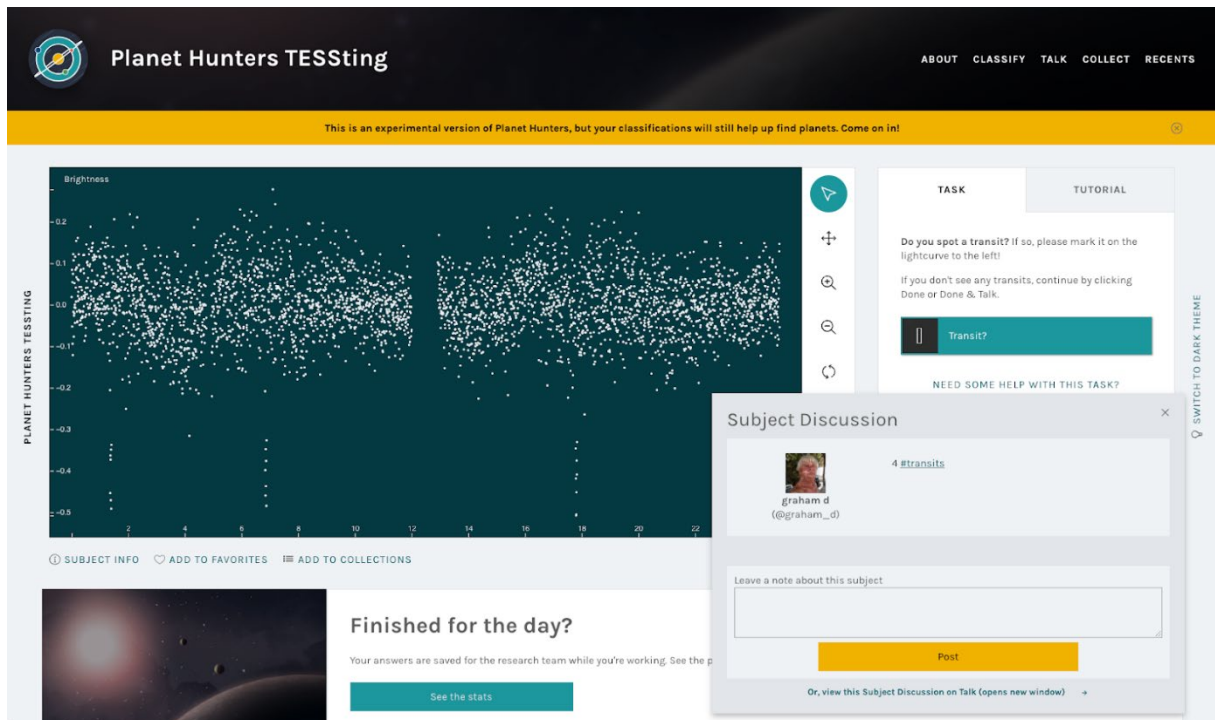
The Zooniverse structure has traditionally supported independent participation in projects only, to avoid bias. Talk is no exception; access to a subject page via 'Done & Talk' is only available after submitting a classification. However, recent Zooniverse-led Digital Humanities research efforts have begun to challenge the assumption that collaboration leads to lower-quality results. In some contexts, it has been shown to produce higher-quality output (Blickhan et al. 2019). In line with these findings, we chose to design and develop a new feature that allows volunteers to view and post Talk comments from the classify page. We called this feature 'QuickTalk'. We initially tested the QuickTalk feature on a test version of an existing project, and it will be featured in the *Communities and Crowds* projects once launched.



A screenshot showing a project classification screen with the QuickTalk icon in the bottom right corner (annotation in red added for illustrative purposes), displaying a subject with no comments.



A screenshot showing a project classification screen with the QuickTalk icon in the bottom right corner (annotation in red added for illustrative purposes), displaying a subject with 1 comment.



A screenshot showing a project classification screen with the QuickTalk dialogue box open, displaying one comment from a volunteer.

2. Community Catalogue

The second effort focused on creating a new space that would allow participants to explore a project's dataset based on not only its institutional metadata, but also the volunteer-generated tags produced through the project message boards. Previous Zooniverse development efforts around exploration and volunteer agency have been limited to institutional archival structures and/or metadata (Blickhan 2021). The Community Catalogue supports exploration and search of not only institutional metadata, but also the tags added to subjects by participants via Talk or QuickTalk.

The main aim of the Community Catalogue is to foreground volunteer-generated information, rather than relegating it to being a secondary source of information about a given dataset. Alongside this development effort, we ask: how might project communities use this tool for exploration as well as adding their own perspectives to the collections at hand? How might the act of tagging differ if a participant knows that the information they are contributing will be immediately available to, and usable by, other participants? How will we need to adjust our expectations of how this tool is used to meet the needs of communities, rather than institutions?

Volunteer-Led Zooniverse Development

With weekly workshops and discussion sessions, the volunteers were given the space and support necessary to lead on the development of the Zooniverse project. In identifying their main goals and objectives, the volunteers emphasised the importance of capturing not only data relating to the digitised archival materials, but also the reactions and experiences of digital volunteers to the subjects and stories illustrated in the archival materials.

Following several exploratory sessions looking at other projects on Zooniverse, as well as the aforementioned paper prototyping session, the volunteers agreed on utilising transcription and classification tasks for the project workflow; the team agreed that these two tasks would gather the most information out of the curated assemblage of materials from the Daily Herald Archives, most of which were photographs. Transcription tasks would require digital volunteers to type out text found on the presented digital object, ranging from descriptive context to handwritten notes. Classification tasks would ask digital volunteers to identify two elements: the presence of edit marks on photographs and any mentioned locations. Together, these tasks would allow for the generation of further data relating to the archival materials that could increase their accessibility and searchability, while also providing the motivation for project participants to engage deeply with these materials during the process of inspection and close reading.

What became apparent in conversations with the volunteers was that alongside gathering as much data as possible from the archival materials, they were very interested in capturing experiences and dialogue from the digital volunteers as well. It was decided that this would be an area where the Talk function (and its new iteration in QuickTalk) would prove useful, with the volunteers themselves acting as both

moderators and active contributors to discussion online in the hopes of provoking deeper engagement with the photographs and documents digitised for this project.

As of the writing of this paper, we are close to launching the latest of the volunteer-created Zooniverse project, 'How Did We Get Here?'. In the meantime, we have ensured that we retain consistent check-ins with our volunteers to ensure that the progression of the project still correlates with their main goals and interests. Embedding an accountable approach has been vital, as it allows us to continuously provide space for the volunteers to take ownership of the creation and development of this work. We aim to continue with this approach moving forwards with the upcoming launch of the project, with the volunteers acting as a review board that we continue to be accountable to in the delivery of the digital project.

Overall, it is apparent that despite unfamiliarity with Zooniverse prior to participation in this project, the volunteers are keen to undertake this next phase of the work online, recognising the potential for international collaboration through digital expansion of access to the Daily Herald Archives, in particular the sharing of knowledge and stories across the wider African Caribbean diasporas.

Audience Research

Throughout this project, SMG's Audience Research team commissioned and carried out focus groups and interviews; this strategic evaluation has provided vital information regarding the volunteers' experience of the project, which will inform future iterations of this new framework of co-creation. Topics of discussion with volunteers included their motivations for participating in the project, their expectations, and whether there were any concerns regarding the current delivery of the project.

Overall, results from the most recent research have indicated two main strands of motivation amongst the volunteers: personal interest and advocacy-based. With regards to the former, volunteers saw the project as a means of exploring their interests in museum work as well as the topics in the archives themselves. With the latter, there was a sense of commitment to the goal of "opening up the archive" through digitisation and crowdsourcing, with some volunteers specifically observing the current disconnect between the museum space and their communities.

Feedback has been mainly positive from the volunteers as well, with most having stated that they felt supported by the rest of the project team and optimistic about achieving the main goals of the project as well as fulfilling their own personal interests in the archive.

As of the writing of this paper, the SMG Audience Research team has held two audience research sessions with the volunteers: once near the start of the project (November-December 2021), and again over halfway through the project (February 2023). We anticipate that a third and final feedback session will take place near the

end of the project to capture volunteer experiences with Zooniverse and digital volunteering, which will be important in evaluating the success of this project overall.

Discussion and conclusions

Ultimately, we hope that the lessons learned from GLAM sector volunteering will help digital crowdsourcing efforts better incorporate community agency earlier on in the design and development process. Creating space for volunteers to participate as project leads means we need to support a broader range of motivations among project leadership. To that end, creating digital platforms and tools that do not inherently prioritise data production is crucial. Additionally, these methods will culminate in a broader examination of how support for volunteer-led projects can enhance traditional project outcomes, such as understanding and interpretation of collections through amplifying community voices.

Next Steps

The workflow outlined in this article is just the first step in creating and sharing a new method for opening up the archive and creating space for volunteers to tell their own stories of heritage collections both in the institution and online. The next steps include creating an iterative space to share the tools and case studies we have discovered, so that they can be applied and tested with new collections, diverse volunteer audiences, and different institutional contexts. An example of one such iterative space that may serve as a model for the *Communities and Crowds* project, is the online Tool Shed developed as part of the *Collecting Otherwise* project at the Nieuwe Instituut (Tool Shed 2023). Utilising a model such as the Tool Shed, we aim to create a space where the tools and workflows we have developed can be adapted, modified or added to by volunteers, heritage professionals or any interested party.

We will be testing and modifying the workflows we have developed with a new set of volunteers at the National Museums Scotland, working to digitise and create a crowdsourcing platform from a large collection of 19th century stereophotographs. We encourage any other institution or individual who is interested in applying this method to their own photographic collections to make this process their own. Together, we hope to ultimately break down some of the barriers to cultural heritage collections and make the archive open for all.

References

Above the Noise, 2019. Accessed at:

<https://www.scienceandmediamuseum.org.uk/what-was-on/above-noise>

App Sheet, 2023. Accessed at: <https://www.appsheet.com/home/apps>

BAATN, 2023. Accessed at: <https://www.baatn.org.uk/>

Benoit III, E., Everleigh, A. (Eds.), 2019. *Participatory Archives: Theory and Practice*. Facet Publishing, London.

Bradford's National Museum Project, 2023. Accessed at:
<https://bradfordsnationalmuseum.org/>

Blickhan, S. 2021. *Engaging Crowds: New Options for Subject Delivery & Interaction*. Zooniverse blog, November 3. Accessed at:
<https://blog.zooniverse.org/2021/11/03/engaging-crowds-new-options-for-subject-delivery-interaction/>

Blickhan, S., et al. 2019. Individual vs. Collaborative Methods of Crowdsourced Transcription. *Journal of Data Mining and Digital Humanities: Special Issue on Collecting, Preserving, and Disseminating Endangered Cultural Heritage for New Understandings through Multilingual Approaches*. Accessed at:
<https://jdmhdh.episciences.org/5759>

Brook, A., Neath, J., 2018. Encounters with Legacy Images: Decolonising and Re-imagining Photographic Evidence from the Colonial Archive. *History of Photography*. 43, 217–238.

Daily Herald Archive, 2022. Accessed at:
<https://www.scienceandmediamuseum.org.uk/objects-and-stories/daily-herald-archive>

Jancovich, L., Stevenson, D., 2023. *Failures in Cultural Participation*, Palgrave Studies in Cultural Participation. Palgrave Macmillan.

Lydon, J., 2017. Democratising the Photographic Archive, in: Reid, K., Paisley, F. (Eds.), *Sources and Methods in Histories of Colonialism Approaching the Imperial Archive*. Routledge, London, pp. 13–31.

Mahr, D., Dickel, S., 2019. Citizen science beyond invited participation: nineteenth century amateur naturalists, epistemic autonomy, and big data approaches avant la lettre. *History and Philosophy of the Life Sciences*, 41, 1–19.

Orr, N., 2006. Museum Volunteering: Heritage as “Serious Leisure.” *International Journal of Heritage Studies* 12, 194–210.

Proctor, A., 2014. *The Whole Picture: The Colonial Story of the Art in our Museums & Why We Need to Talk About it*. Cassell, London.

Ridge, M. (Ed.), 2014. *Crowdsourcing our Cultural Heritage*. Ashgate, Surrey, UK.

Ridge, Mia, et al. 2021. ‘1. Introduction and Colophon’, in *The Collective Wisdom Handbook: Perspectives on Crowdsourcing in Cultural Heritage - Community Review Version, Community Review edition*.

Simon, N., 2010. *The Participatory Museum*. Museum 2.0, Santa Cruz.

Science Museum Group (2020) *Inspiring Futures*. Accessed at <https://www.sciencemuseumgroup.org.uk/wp-content/uploads/2020/05/SMG-Inspiring-Futures-May-2020.pdf>

Tauginiené, L., et. al., 2020. Citizen science in the social sciences and humanities: the power of interdisciplinarity. *Palgrave Communications*, 6, 1–11.

Tool Shed, 2023. Accessed at: <https://nieuweinstituut.nl/en/projects/collecting-otherwise/toolshed>

Van Broekhoven, L.N.K., 2018. Calibrating relevance at the Pitt Rivers Museum, in: Pellew, J., Goldman, L. (Eds.), *Dethroning Historical Reputations: Universities, Museums and the Commemoration of Benefactors*. University of London Press, pp. 65–80.