OXFORD

Enriching Cultural Heritage Communities: New Tools and Technologies

Alan Dix[1,2], Elizabeth Jones[1,2,3], Rachel Cowgill[4], Charlotte Armstrong[5], Rupert Ridgewell[6], Michael Twidale[7], J. Stephen Downie[7],

Maureen Reagan[7], Christina Bashford[7], David Bainbridge[8], Carys-Ann Neads[3] and Vince Davies[3]

[1]Swansea University, Swansea, Wales, UK
[2]Cardiff Metropolitan University, Cardiff, Wales, UK
[3]Troedrhiwfuwch Community, UK
[4]University of York, UK
[5]the Valuable 500, UK
[6]British Library, London, UK
[7]University of Illinois Urbana-Champaign, USA
[8]University of Waikato, Hamilton, NZ

This paper explores ways in which scholarly skill and expertise might be embodied in tools and sustainable practices that enable communities to create and manage their own digital archives. We focus particularly on tools and practices related to the recording and annotation of digitized materials. The paper is based on co-production practice in two very different kinds of community. Although the communities are different we find that tools designed for a specific community are valuable for others, thus offering the promise of general tools to support community-centred digitization and potentially also traditional archival practice.

RESEARCH HIGHLIGHTS

- A co-design study with a diaspora community developed an oral-history application to help connect and enrich their archives and memories.
- A second study explored collaborative approaches to assembling and digitising runs of concert programmes and other data-rich musical ephemera.
- These consider ways of making community archives accessible for research and engagement for all.
- Bespoke tools developed for one setting were also useful to the other.

---

1 Introduction

It is often said that Covid-19 has awakened many to the importance of community. However, it also seems that communities are under threat, their sense of identity and belonging drowned in the homogenization of global media and rootlessness of modern living. The heritage, culture and history of communities is one of the things that nurtures this sense of belonging; and the importance of cultural heritage is well recognized, both amongst researchers (e.g. Giglitto *et al.*, 2019) and in the UNESCO (2003) "Convention for the Safeguarding of the Intangible Cultural Heritage". Yet heritage is also precarious. When flood or fire destroy some part of a national museum or art gallery it makes headline news, whereas shoe boxes of memorabilia or old papers are discarded or lost every day, as people move, die or downsize.

One of the bulwarks against loss of large, institutionally supported collections is digitization, which also opens the archive to more widespread scholarly study of materials and public dissemination. Furthermore, in the internet age, digital resources offer greater visibility and thus influence. The role of the latter was particularly crucial during periods of Covid lockdown (Vayanou *et al.*, 2020), but also has the potential to allow materials to be re-presented in ways that reach audiences who would not usually visit cultural institutions. Digitization can also bring scattered archives together, virtually, without requiring ownership of the objects themselves to be relinquished by a donor.

Can these benefits of digitization be harnessed for small communities, offering them the means to preserve, explore and publicize their own heritage and stories? More crucially, can we *democratize digitization*—make it available, not simply when a team of university researchers parachute in to offer expertise and resources that are necessarily limited in time and scope, but embodying that skill and expertise in tools and sustainable practices that enable communities to manage their own digital archives?

Of course, the majority of the authors are just such a team of university researchers. Some of the challenge of this kind of work is mutually valuing the variety of different knowledge, skills, situated understanding, and experience that we all bring, while respecting the differing needs of stakeholders, including those in the communities, academia and funding bodies. There are few easy answers beyond maintaining an openness to others and readiness to reflect (Avram *et al.*, 2020, Giglitto *et al.*, 2018).

In this paper we explore some of the questions around these issues and present early prototypes that we hope will be valuable across different kinds of communities and settings. In particular, we will concentrate on tools and practices related to the recording and annotation of digitised materials—that is, the creation and management of digital community archives. We will also see that the boundaries between collection, curation and communication are far more fluid in community-centred digitization than in traditional archival practice... even though the latter is itself at a point of flux (Dix *et al.*, 2014b, Hoyle, 2022).

We focus on two very different kinds of community, both of which are *pseudo-geographic* in that they have elements of physical locality, but include members not defined simply by where they live. One is a small village, Troedrhiwfuwch, in the Welsh (ex)coal-mining valleys, that was evacuated in the 1980s due to concerns about potential landslips, and physically demolished, save for a building or two. A sense of spirit about the community lives on, however, in those that lived there and their descendants. The others are a group of local music societies in Belfast, Huddersfield and York in the UK, all of which originated in a widespread post-First-World-War initiative to use music to rebuild a sense of international connection.

Note that these communities were not carefully chosen for their similarities and differences, but rather engagement began through coincidences and personal contacts: in the case of Troedrhiwfuwch through a researcher who was also a member of the community, and in the case of the local musical societies because one of the researchers lives and works in two of the three locales. They arose in different projects, with their only initial linking point one researcher who was involved in both projects, and a common desire in each community to gather and share their heritage.

While the social demographics and reasons for existing are very different, we will see that there are commonalities—principally, prototypes that were designed for each have value for the other. This suggests that bespoke development and rich co-design for specific communities can lead to tools and processes useful to many. It has been noted by others that participatory approaches are, by their nature, 'unique to each project' and thus it can be hard to leave a broader legacy beyond the particular 'project-specific endeavours' (Avram *et al.* (2020), p.255). In contrast, whilst there is, of course, methodological learning from each instance, in addition we shall see the potential for technical legacy.

In the next section we will review some of the conceptualizations of the word 'community', which is critical in so many disciplines, from human geography and social science to health, besides heritage. We then look at the two communities we are studying: Troedrhiwfuwch (Section 3) and the former regional branches of the organization founded as the 'British Music Society' in 1918 (Section 4). For each we will first describe the community, the engagement between researchers and community members, and initial concepts and themes emerging from them. We will then look at a prototype designed for the community: TalkOver for Troedrhiwfuwch, and OcrMarkup for the music societies. After describing each community and its prototype, we will look, in Section 5, at what happened when the communities were exposed to the prototypes for the other community, and consider lessons we can take away from this.

2 Dimensions of Community

Community is a word we all recognise and yet almost certainly all understand in different ways. Most readers of this paper will be academics and in parallel be part of: a local community around

their home; maybe a separate 'home' community where they were brought up; a university or departmental community, including academics, administrators and students; and a professional community, for example as HCI researchers.

The AHRC Connected Communities programme in the UK (itself a community of practice of researchers of 'community') produced a number of detailed reviews and commentaries, which together capture some of the complexities of community (Crow & Mah, 2012, Studdert & Walkerdine, 2016). This includes the way 'community' emerged as a subject of study in the 19th Century, largely in response to something being lost (Walkerdine & Studdert, 2012), and highlights that 'community', despite a large literature and being the focus of many government initiatives, is still often poorly defined—a 'spray-on term' (Walkerdine & Studdert, 2012), or 'slippery concept' (Craig & Mayo, 2011) subject to 'disciplinary confusion' (Studdert & Walkerdine, 2016).

The most obvious concept of community is geographic—people in a village, town, or urban neighbourhood, the idea of one's own *campanilismo* (bell tower) in Italy, or *milltir sgwâr* (square mile) in Wales. However, researchers in human–computer interaction will also be familiar with the anthropological concept of 'communities of practice' (Lave & Wenger, 1991, Wenger, 1999), which are often linked to professions, or other forms of interest group.

This distinction, geographic vs. interest-based, is fluid and many communities are *pseudo-geographic*: they may be associated with a specific place, albeit not necessarily living in that place (e.g. university alumni); or they may reside in a smaller/larger space, but be based around interests or characteristics in common that are not shared by everyone in the region (e.g. religious or ethnic communities within an area, or chambers of commerce). The communities we will describe below are both pseudo-geographic—one dispersed, but linked by a single common physical origin, the other based around common musical interest within a wide geographic area.

As well as these dimensions of place and interest (Willmott, 1986), many conceptualizations look more at what a community does or how it is experienced, including a sense of identity (Willmott, 1986); imagined affinity (Anderson, 1983); a matter of feeling (Cohen, 1985), connection, difference, boundaries and development (Crow & Mah, 2012); or the action of communing, relationality and sociality (Studdert & Walkerdine, 2016). According to one World Health Organization definition (Nutbeam & Kickbusch, 1998, p. 354): "*Members of a community gain their personal and social identity by sharing common beliefs, values and norms which have been developed by the community in the past and may be modified in the future*". Further, Kay Kaufman Shelemay (Shelemay, 2011, pp. 349–350) emphasises the importance of ritual and repetition in community formation, arguing in relation to music that performance and transmission play not just a symbolic role, but a dynamic one, "*as an integral part of processes that [...] help generate, shape, and sustain new collectivities*".

These characteristics of community bridge geographic and thematic dimensions and emphasise the shared aspects of communities of many kinds. It is therefore not so surprising that we shall find that tools created for one kind of community end up being applicable to others.

3 People of a lost land
3.1 Context

Troedrhiwfuwch was founded as a small coal-mining village nestling on the eastern slopes of the Rhymney Valley in South Wales. From 94 households, 110 young men left for the First World

FIGURE 1. Troedrhiwfuwch before and after evacuation; just one house (centre) and the former post office remain (Source: Troedrhiwfuwch community archive).

War, 21 of whom never returned. This was one of the greatest concentrations of war-service enlistment in the country for the size of the small community, which totalled 600—a commitment and sacrifice recognised by King Edward VIII in 1936. Then in 1976, the village was condemned. In 1966, 28 adults and 116 children lost their lives in Aberfan, another mining community, when a rain-soaked coal tip, the discarded rocks and coal dust from deep mining, slid down the mountainside and buried the village school. In the aftermath, surveys assessed the stability of other mountain sides and coal tips across the area. The mountain above Troedrhiwfuwch was deemed at risk, and, over a number of years, the people of the village were rehoused. Most of the village structure was demolished by 1985. Today, only two houses, and a war memorial and garden remain as a sign of the place that once was (Figure 1).

The diaspora of Troedrhiwfuwch, or 'Troedy', as it is known locally, has not forgotten its past. Each year on Armistice Sunday, a group congregates at the War Memorial for an act of remembrance and there is an active Facebook memories group. A smaller group is also active, gathering photographs and documents from local people and scouring national archives for material connected to the village. This includes a digital archive of more than 1,400 items, at the last count, and extensive paper material. A particular focus has been on the First World War, especially following the centenary events of 2014–2018, and given the War Memorial and the adjacent Memorial Garden (on the site of the demolished church) are some of the few remaining signs on the ground.

This diasporic community is not just the old who lived their lives there, although some are in this category. Many only know of the village through trips as children, when parents and grandparents would point to a patch of grass and tell them stories of the place where an aunt or cousin once lived. Some lived in or visited the village as a small child while it still stood, but one of the most active members of the history group was born well after the long terraces of houses, which once lined the roadside, were demolished.

3.2 Engagement

One of the authors of this paper works for a university as well as being a family member of the Troedrhiwfuwch community. She acted as the first point of contact for the project. Since March 2021, a small group of academic researchers and community members have met, largely informally, around a dozen times. This has been mostly using video conferencing, but there have also been several site visits, albeit limited ones in the early days due to Covid. The latter included walking the ground of the village itself, and also visiting a church in a neighbouring village where the interior furnishings of the demolished Troedrhiwfuwch Church have been used to create a small side-chapel, forming a compact reproduction.

As appropriate to any co-production exercise, the team wanted to embed the principles of equality, diversity, accessibility and reciprocity in putting co-production into action (Social Care Institute for Excellence, 2015), and there was a period of mutual enculturation. On the one side, the non-Troedrhiwfuwch academics built an understanding of what it means to be part of the community. This was accomplished principally through *story-telling*, often focused around digital artefacts, or walking the ground itself. On the other side, the community members built an understanding of the potential of digital technology to help them preserve, organise and disseminate their heritage materials. This was facilitated by the production of early envisionments using PowerPoint scenarios and paper-and-card low-fidelity prototypes. These effectively acted as a form of technology probe (Hutchinson, 2003) allowing the participants to see the potential of available technology without committing to a particular design path.

3.3 Emerging concepts

One way to view engagement between university researchers and community members would be as an expert–amateur or expert–end-user conversation, based on mutual respect but with different roles. There is truth in this; however it misses the rich and diverse **expertise of the community** members themselves. There are obvious elements to this expertise: personal knowledge of events through direct experience or conversations with others—connection points into human networks and understanding of the needs and aspirations of the community—but this is only a part of the story. The Troedrhiwfuwch volunteer archivists have a knowledge of historic sources such as military records, genealogical resources and census reports. This facility with primary resources is complemented by a synthesised knowledge of the historic relations between people and events, similar to that which the (non-historian) academic members of the team have observed in their academic colleagues' historical knowledge. This is not to equate academic and community historical expertise and approaches, but to problematize words such as 'amateur' and 'expert' (Armstrong *et al.*, 2023).

One of the key differences is that community history is often **intimately connected to family history**. The people in a photo are not simply objects of study, but great-aunts and grandparents,

FIGURE 2. The Troedrhiwfuwch diaspora—many of the community still live close to Troedrhiwfuwch, but some are scattered across the UK and the world.

with stories that are part of one's own story. Equally, these ***personal stories are often universal stories***, and (for those outside of the community) the stories of individuals to whom one has no personal connection do not merely fascinate as stories, but can parallel one's own experience—lessons not lost on the producers of popular TV family-history programmes.

We have noted the extensive nature of the existing digital archive including photographs, documents, census records. Whilst the individual items are preserved and organised, the meta-data—the knowledge of what things are and how they relate to one another—is largely in the heads of the community archivists. This includes the provenance of items—who donated a photograph or pamphlet, and from which website or military archive an item was downloaded. This is important from a scholarly viewpoint, but also practically – for example, if items are presented externally on a community website, are there intellectual property (IP) restrictions on images? Filesystem design has hardly changed since the 1970s—each file is isolated and related to others only by their location in the folder/directory hierarchy. Archivists, both professional and lay, need better ways to ***annotate and connect***.

Expert knowledge is often tacit—only brought to bear in particular circumstances and contexts. This is equally true of community knowledge—***people, places and artefacts*** elicit knowledge and stories. One example of this was seen while walking the ground of the village. The precise position on the ground of demolished houses was often half-guessed in relation to natural outcrops of rock. Then one of the community members said, 'my house was here'. The house had gone, but the drain in the road had lain by the outer corner of the house and the drain remained.

There is a ***fragility and precarity*** to these memories. This is true of the personal memories of ageing people, but also for physical artefacts. Troedrhiwfuwch emphasises that even buildings and solid rock may shift or fall. Between memories and masonry are many photographs, small items and documents that live on mantelpieces or in attics. When a person dies, not only are their memories lost, but these objects, embodying community heritage as well as personal significance, may end up on the fire, or in a junk shop or skip. This precarity has been noted elsewhere; for example, Giglitto *et al.* (2019) report a concerning 'abandonment of storytelling' amongst Bedouin due to the nature of contemporary urban life.

Within the research community there is an increasing push towards open resources. However, while the community archive has been widely shared internally, there are clear ***limits to openness***, boundaries as to what should or should not be made available openly, particularly on the web. This is partly due to the fact that some material is derived from non-open sources, such as subscription web services. Moreover, even material in the public domain may not be suitable for sharing – for example, archival newspaper reports of potentially embarrassing court cases that it would be insensitive to place in an open repository.

3.4 Prototype: TalkOver—capturing stories about photographs

TalkOver is an experimental web app that makes it easy to record stories about pictures. It can be used for gathering oral history about old photographs or documents, or for any application where you want to produce narratives about images.

TalkOver was not amongst the early envisionments used during the co-production process. Instead, the need for it arose more gradually out of experiences during meetings between the Troedrhiwfuwch community and researchers. The extensive archive of photographs and documents is impressive in itself. However, as soon as any one of the photos is opened, community members start to tell stories: some about past relatives that they were told as children, some from research they have done in other archives or war records. The details that make the photographs come to life and connect them together are in the heads and memories of the community, but not recorded in their digital archive.

Narrative and storytelling have always been an essential part of community history. The cites examples from as far back as the 8th Century, and Sharpless (2008) looks back to Heroditus, in the 5th Century BCE. This accelerated in the 19th Century, especially in relation to folk tales and songs. However, the emergence of audio recording, and especially magnetic-tape recording, created the modern field of oral history. Digital technology has further transformed the collection and curation of audio material (Lambert & Frisch, 2019), for example it is now possible to geo-code stories whilst walking so that they are connected with particular locations (Zembrzycki, 2013). In presentations of oral history for public access, the spoken word is often illustrated in professionally edited multi-media presentations, with the voice overlaying still images. Based on experiences during the participatory sessions, it became clear that something similar was needed, but with the ease of pointing at people in a photograph as one does when sitting side-by-side with someone.

TalkOver addresses this not just by recording the speaker's voice, but also by allowing the person being recorded to point at a digital image, using either their finger on a tablet or mouse on a laptop screen. As the user touches the picture, a small halo temporarily appears at the point they touched, as feedback (see
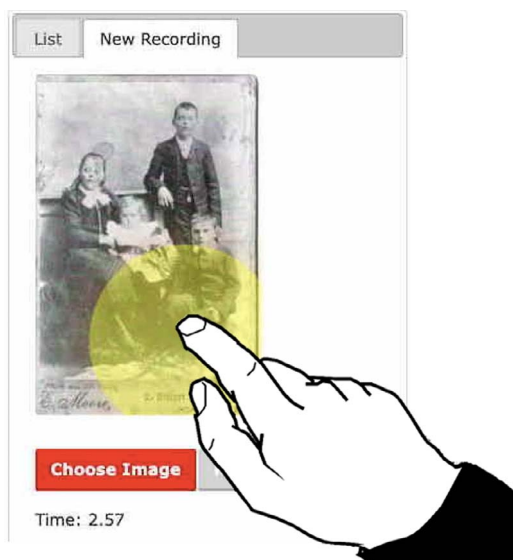
FIGURE 3. TalkOver recording in progress.

FIGURE 4. Pictorial list of TalkOver recordings.

Figure 3). The locations highlighted on the images in this way are recorded along with their time-stamps. The audio and marks are stored alongside the image and can then be replayed. This creates a rich playback akin to a crafted multi-media presentation, but with the immediacy of a side-by-side telling. As the work was performed during Covid lockdown, this was especially poignant.

As well as offering an enriched form of collection, the marks associate areas of the image with points in the story. If faces or objects in the images are also indexed, automatically or by hand, with people and themes, then this offers the potential for interlinking semantic annotations and continuous media.

There is an art to interviewing for oral history, and the system does not replace that. However, the act of talking about something is often very natural and thus offers a way for less skilled interviewers to collect oral history, as well as providing an additional tool for the professional oral historian. In particular, a likely scenario of ongoing use is inter-generational, where younger members of families or the community, such as school-children, use TalkOver in combination with other tools to collect reminiscences from grand-parents and other older members of the community.

3.5 Under the Hood

TalkOver is built as a standalone web app—that is, all processing and storage are local to the user's machine. This allows sharing of usable prototypes, without complex installation and without the need for extensive cloud or server infrastructure.

New images can be added by drag-and-drop or the file chooser, using standard cross-browser W3C file APIs. WebAudioRecorder.js is used for audio capture, which is built on the W3C WebAudio API. This performs all recording and encoding in Web Workers in the browser, meaning that no external transcoding is needed. The audio is stored in the browser's IndexDB store, which can accept large media data and provides persistent local storage. This is used to store both the raw media (images and audio) and the data structures describing the user's pointing actions (essentially time-stamped coordinates). A simple pictorial grid is used to select previous recordings (Figure 4).

Given the use of web technology, the import/export format for backing up and sharing TalkOver recordings is simply an HTML file (see Figure 5). This includes the complete image and audio media base64-encoded in JavaScript variables, as well as further meta-information in sections demarcated by easily identifiable comments. These TalkOver HTML archives can be loaded back into the TalkOver application, which parses the HTML and extracts the media variables. A Globally Unique Identifier (GUID) is generated for each new TalkOver recording and stored in the HTML archive format, so that if a backup is reloaded or a shared recording loaded twice it can be connected to the original recording.

The HTML content is minimal, but includes a link to a single JavaScript bootstrap file, which allows smaller recordings to be opened by double-clicking the HTML file without explicitly importing. This is similar to the self-describing '#!' prefix for running script files in Unix. As they encapsulate all the media, these HTML archive files are large (around 120Mb for a ten-minute recording), but recordings of up to two minutes have been 'click opened' in Safari and up to five minutes in Chrome (both in MacOs). We have not yet hit the limit for import/export sizes, and so, as it is intended to be for relatively short recordings, this format seems sufficient for the purpose.

4 Regional music societies
4.1 Context

Among the music clubs and societies active in the first decades of the 20th Century, the British Music Society (BMS) stands out for its ambition, reach and impact. It was established in late 1918 to restore international collaboration and exchange between British and overseas musicians after the twin catastrophes of the Great War and Spanish Influenza, and to empower amateur musicians and music-lovers in organizing and promoting their own concert series, providing mostly professional classical musicians with paid engagements and infrastructure to help rebuild careers and establish new ones. The BMS was formed by the progressive musical author, educator and organist Arthur Eaglefield Hull

```
1   <!DOCTYPE html>
2   <html>
3   <!-- @@wrappeddata      -->
4   <!-- @version: 0.1      -->
5   <!-- @type: talkover    -->
6   <head>
7       <meta charset="UTF-8">
8       <title>Talkover</title>
9       <meta name="viewport" content="width=device-width, initial-scale=1.0">
10  </head>
11  <body>
12      <h1>TalkOver - Playback</h1>
13      <div id="playback"></div>
14      <br clear="both">
15      <script>
16          // @!meta
17          var meta_info = { "type":"talkover" };
18          // @!end meta
19          // @!data
20          var talkover_info = {"version":"0.1","creation_time":"2022-01-16T19:35:19.113Z","au
21          var image_url = "http://whereweare.org/test/record/images/Hillman-family.jpg";
22          var audio_url = "data:audio/wav;base64,UklGRiTACQBXQVZFZm10IBAAAAABAAIARKwAABCxAgAE
23          // @!end data
24      </script>
25      <script src="https://datatodata.org/talkover/demo/v001/js/bootstrap_v0.js"></script>
26  </body>
27  </html>
```

FIGURE 5. TalkOver export format as HTML; note the base64-encoded `audio_url` is typically between 5 and 50 million characters long.

(1876–1928), with chapters[1] opening in towns and cities throughout the UK and beyond (Cowgill, 2018). Although the BMS as an organization was wound up in 1933–34, some societies descended from these chapters continued, flourished, and remain active today (note: the British Music Society founded in 1979 is an entirely separate organization). While they may have limited knowledge of their shared origins in Hull's BMS, they have amassed substantial archives over the past century that shed significant light on the rich history of this extraordinary initiative and the broader role of music in regional community life.

BMS chapters were also established overseas, mostly along colonial pathways, as in the case of the Bangalore branch of the BMS in India[2] Cousins (1935), raising significant questions about the 'performance' and meaning of Britishness and internationalism in these contexts (Cowgill, 2022). The BMS would also become the launch-site of the British Section of the International Society for Contemporary Music (ISCM) in 1922–23, a relationship still not fully understood (Arrandale, 2023, Cowgill, 2022, Kelly, 2023, Masters, 2021).

Designed to locate, digitise, consolidate, enrich and interrogate archives such as these, The Internet of Musical Events: Digital Scholarship, Community, and the Archiving of Performance (Inter-MusE) was established as a two-year project (2021–23) with funding from the AHRC's UK-US New Directions for Digital Scholarship in Cultural Institutions scheme (Ref. AH/V009664/1). InterMusE has brought together a team of scholars from humanities and computing backgrounds to work with three former chapters of the BMS: the Belfast Music Society (BeMS), British Music Society of York (BMSY) and Huddersfield Music Society (HMS). These institutions are eager to take stock of their histories and document their collections, and InterMusE has been working with them as a case study to capture and link different forms of data relating to historical musical events with a view to creating a dynamic,

open-access digital archive of musical ephemera. Befitting the international aspirations of the founders of the BMS, Figure 6 shows how the project partners and source archives are spread across the globe.

The collections of the BeMS, BMSY and HMS comprise diverse material types, from concert programmes, season prospectuses, and other performance ephemera, to newspaper reviews and administrative records Armstrong *et al.* (2023), Bainbridge *et al.* (2023) (Figure 7 shows an item from the HMS collection). In each case, some physical materials are stored in local archives or libraries, while others are kept in society offices and private homes. As such, the materials have undergone varying degrees of cataloguing, digitization and preservation. Each society has representative members, volunteers or employees, who have taken a keen interest in its archival collection. Drawing on a range of professional, self-taught and instinctual knowledge, these representatives—the custodians of the collections—have taken steps to ensure the preservation of their society's archival materials for future generations. By working with them to capture and link the data from these materials, as part of a unified digital archive, we aim to improve access to the archival collections and empower society members to explore and engage with their rich histories, including relationships between local branches and centres 'on the ground', as it were, and the umbrella organization (BMS) under which they operated. This will include opportunities currently opening up to link with the archives of BMS chapters in New Zealand (Whanganui), and Australia (Sydney and Melbourne) (Kirby, 2023ab). We are also exploring ways in which the expertise of these community members can be used to enrich the historical records of these societies incrementally. The digitised materials will be enhanced with item descriptions and transcriptions, personal recollections and oral histories. Isolated or short runs of documents from other UK BMS branches are continuing to surface in libraries and archives—these include concert programmes from London (Marylebone, Hendon & Golders Green, London Contemporary Music Centre), Birmingham, Blackpool, Bournemouth, Bradford, Leeds, Liverpool, Manchester and Newcastle, to date, and are being added to the digital archive.
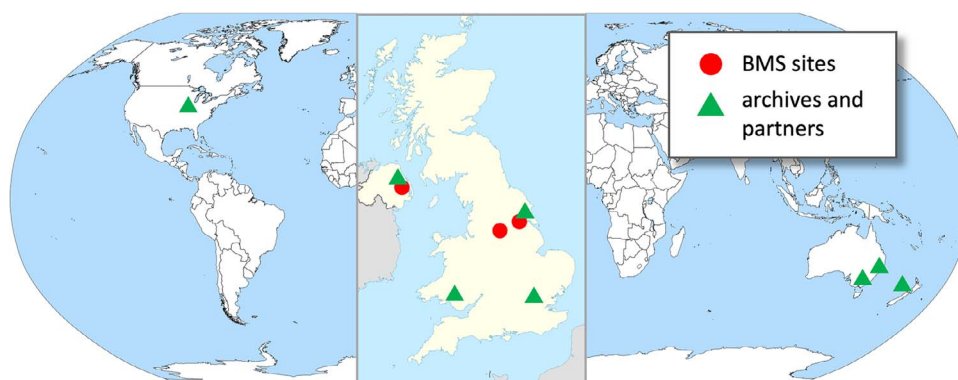
---

[1] The BMS divided its local societies into 'branches' and 'centres', primarily on the basis of size, so we have used 'chapters' as a generic term, here, to avoid confusion.

[2] Commemorative copper plaque (1927), privately owned, India, inscription: "Presented to O. Schmidt ESQ. by the Bangalore Branch of the British Music Society in Commemoration of the Beethoven Centenary 26th March 1927. —— Opened by His Highness Maharaja Bahadur Shri Harisingh Bahadur Maharaja of Jammu and Kashmir State in March 1927 in the first year [of his reign]."

FIGURE 6. InterMusE partners and sites of continuing BMS chapters and community archives.
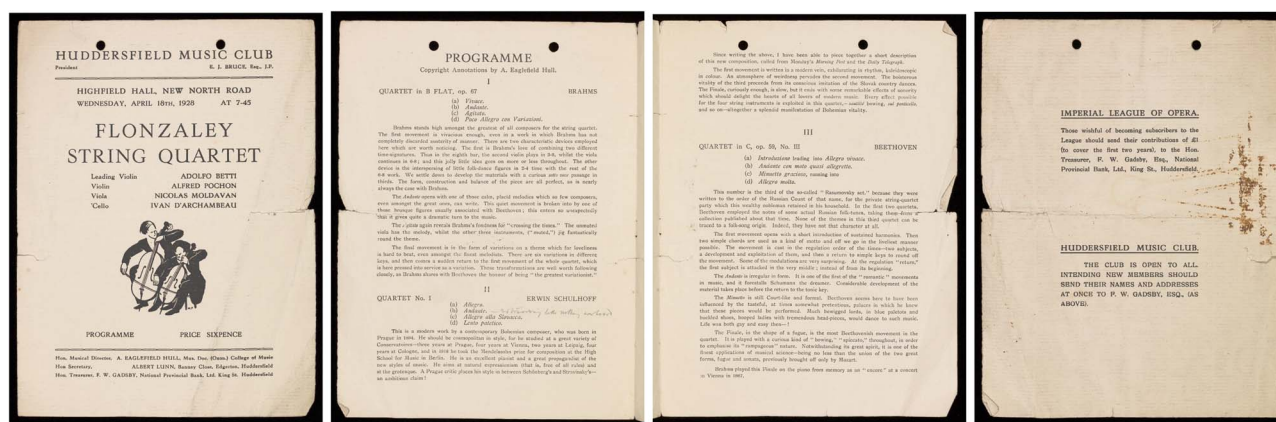

FIGURE 7. Example of a 1928 concert programme from Huddersfield Musical Club, the name by which the Huddersfield chapter of the BMS was known at that time (Source: HMS archive).

4.2 Engagement

From university-based researchers, archivists and programmers, to citizen researchers, amateur musicians, music-lovers and audience-members, InterMusE brings together a range of different stakeholders. The project places a strong emphasis on collaboration and co-production with these societies and their communities, and resists privileging any one stakeholder group over any other. To ensure that the digital archive produced is both a valuable research resource and fit-for-purpose for the societies, the approach has been shaped by a desire to design and create a digital archive with (rather than for) the societies and their communities. Of course, the fact that the work was funded by the AHRC means that novel research had to be a central aspect; however, this is interpreted widely, and the co-production focus of the project was not only accepted, but welcomed by the funding body.

One of the first steps was to take stock of the current collection and preservation activities in each society and understand various stakeholders' visions for the project. These collection assessments were conducted over Zoom in April 2021 as informal, unstructured interviews. This kind of informal interaction proved effective in establishing a foundation for trust and reciprocal exchange between the project investigators and citizen groups (Armstrong *et al.*, 2023). In July 2021, a second set of group information sessions, also conducted over Zoom, provided a forum for society members to share their thoughts on the project and voice any questions or areas of concern.

4.3 Emerging concepts

Several of the themes that arose in Troedrhiwfuwch have parallels in the music societies.

The ***expertise of the communities*** was again very evident. Some of this is in terms of skills and experience brought into their roles; for example, the music-society committees include several members who have retired from senior roles in public service and industry, including arts management and creative careers. In addition, one member (a former professional librarian) has developed a complete database of concerts including itemization of the programmes.

The interweaving of ***community with personal and family history*** is also evident, although in a different way. Committee members are often long standing, so when looking through old committee minutes or concert programmes they see names of current and past friends and family. In addition, the concert venues mentioned in early 20th-century programmes are typically in local places, and in many cases are still standing and may even be active or recent venues. That is, ***people, places and artefacts*** elicit knowledge and stories in a very similar way.

Although elites in London may have refered somewhat condescendingly to 'the provinces' in publications, people with national and international reputations often travelled to places like Huddersfield to perform, sometimes as part of an extended tour (Cowgill & Holman, 2007). Early investigations in and beyond the InterMusE archive by humanities students in Illinois have uncovered many connections between performers in the UK societies

and the classical-music scenes of the USA and continental Europe, speaking to the **universal** significance of local history within the global community of interest—exactly Arthur Eaglefield Hull's vision. By connecting information in the digitised concert programmes to other databases we can see richer connections with larger social and political events, such as the Russian Revolution and the disintegration of the Austro-Hungarian Empire, and how that impacted upon who performed what in British towns and cities in the 1920s.

Issues of **fragility and precarity** are also common in discussions. When a member of one of the societies, who had an extensive collection, died, the documents could easily have been lost; but their spouse knew about them and the passion for preservation that lay behind them, and was able to pass them to a current committee member. It was evident, however, that this was a moment when crucial records might have disappeared for ever. Concert programmes are particularly undervalued as historical sources among musical ephemera, and frequently disposed of or recycled after the events they were produced for have passed, including in house clearances (Armstrong *et al.*, 2021).

In the Troedrhiwfuwch archive, many of the textual items (e.g. war records) are already digital, and many of the more internal community artefacts are photographic and visual. In contrast, the music societies have large paper repositories of largely textual and formalised content, such as concert programmes, reviews and meeting minutes. The immediate need is to **digitize and then extract relatively structured information** from them. That is, while the need to **annotate and connect** is present, the material is of a more structured form, even though the individual formatting of that information differs from programme to programme.

As with the Troedrhiwfuwch archive, there are **limits to openness** within the BMS archive, principally related to intellectual property and permissions. In particular, the contributors of programme notes and reviews of musical pieces may not be available to give permission for their use in other media, notably the web. It is hoped that the formal archive items (programmes, season prospectuses, minutes, newspaper cuttings) may be augmented over time with textual and oral reminiscences about the material. In this way, as more personal material is added, issues such as content moderation, restricted access and time-locked material may need to be considered. Data protection and consent, of course, are required from the off.

4.4 Prototype: OcrMarkup—from text to meaning

This envisionment prototype was created to show how OCR can be used to help add semantic markup to scanned documents. This is specifically for situations where a level of expert judgement is important… that is, where a fully automated solution is not appropriate, but we still wish to make the most of what the computer can do to help.

This prototype arose directly from early discussions in the InterMusE project where we are working with concert programmes. Commonly the output of OCR is a continuous text, sometimes with attempts to deal with common forms of document structure, such as columns. The text versions of documents in Project Gutenberg or HathiTrust archives are good examples of this. This works well for linear text, such as a novel, but less so for structured documents.

In previous projects we had created digital versions of two 19th-century catalogue-style documents, the *British Musical Biography* (Brown & Stratton, 1897) and *Gazetteer of Scotland* (Wilson, 1882). These were semi-structured, although care was still needed to

FIGURE 8. OcrMarkup showing areas marked on screen and annotation fields.

identify entry headings (personal names or place names) semi-automatically, for example, using all-caps.

Concert programmes are far more complex, with multiple sections for performers, dates and times, pieces played, etc. (see Figure 8, left). Complex many-to-one and one-to-many relationships are communicated visually (and via conventions learned through familiarity) in a programme by the layout and different sizes and typefaces used for the text, such as the movements of a string quartet, the composer or arranger, and performers playing, usually in a particular named ensemble (see Figures 7 and 8). It is important to extract this rich information, but there are variations between programmes and a substantial portion of the text consists of personal names and titles of pieces (in a variety of languages), making automatic processing difficult. For example, off-the-shelf OCR might take a column of performer names and concatenate it into a single unpunctuated paragraph:

```
ADOLFO BETTI ALFRED POCHON NICOLAS MOLDAVAN IVAN
D'ARCHAMBEAU
```

The variety of concert-programme structures means that **human-intensive intervention** is essential in order to extract meaningful semantics. Happily, for community-based digitization that human-intensive intervention is possible (Dix *et al.*, 2019), although we also want to make as much use as possible of OCR in order to make the human task as fluid as possible.

While the final version of OCR is often a linear text, earlier stages of the OCR pipeline retain the precise location on the page of each character, word or phrase. Google Vision API was used initially for OCR extraction in this project, but the current prototype uses Tesseract.js if there is no existing markup. The latter occasionally misses words that are recognised by the Google cloud service; but the differences are marginal and for community use the advantages of open source and a free-at-point-of-use service outweigh the slightly better quality of Google Vision. In later versions we plan to allow configuration of OCR services, including use of OCR embedded in PDF when available.

The OcrMarkup prototype allows the user to select and name areas of the image and automatically extracts the OCR text for the region. Figure 8 shows this in action. The user has dragged out a series of areas in the image and then for each region, as it is selected, the text for that region is placed in a corresponding area in the right-hand column. The user has then labelled these areas 'venue', 'date', 'time', 'title', and is in the process of typing 'performers' for the most recently identified section. If the user resizes the section on the image, the text in the named annotation is automatically adjusted.

On its own, OCR is useful to allow free-text searching of large digitised collections. It is also possible to automatically identify common types of data, such as dates or personal names. However, when a human looks at a document they can identify more detailed and specific areas, such as the title of a concert, or who was performing, creating a rich semantics for each document. While this human-in-the-loop identification of areas is a simple technique, the only other system of which we are aware offering such a facility is Lace0.5 (Robertson, 2021). Due to a difference in the use case—semantic markup of the Open Greek and Latin corpus—it adopts a fixed vocabulary for marked sections rather than the open annotations allowed in OcrMarkup.[3]

Lambert & Frisch (2019) describe their transition from linear models of content curation to a hub model, where a core of raw data (e.g. recordings or photographs) gives rise to numerous smaller or larger collections of 'cooked' data, interpreted and annotated by different tools for different purposes and audiences. Our own work also emphasizes these more incremental approaches, layering different interpretations and processing, automatic and human, by scholar or community (Armstrong *et al.*, 2021, Dix *et al.*, 2014b).

OcrMarkup fits into this broad process. Annotations are added incrementally based on the purpose and goals of the user. For example, when a programme is first scanned, a community archivist may simply want to annotate key features, such as the date, venue and title of the concert, in order to create a bare-bones listing of events. Later another community member might be looking for references to a particular family of musicians, using free-text search to find candidate documents and then marking up relevant parts. Each person's efforts add to an evolving semantically annotated digital archive.

4.5 Under the Hood

Like TalkOver, OcrMarkup is built as a standalone web app for ease of distribution.

The core application consists of four main elements, each relatively simple in their own right:

OcrManager – A wrapper class for OCR text.
ImageMarker – Managing the selection of areas on the image.
FieldManager – Managing the right-hand panel where fields are named and edited.
AnnotationArea – A coordinating agent, linking image areas and field definitions using Observer-pattern events, and also managing the interface with persistent storage and import/export.

As noted, the first prototype used Google Vision API, but the current version uses Tesseract.js as this executes within the browser (asynchronously as a Web Worker). The `OcrManager`, however, makes the annotation code independent of the choice of OCR engine and, where present, Google Vision OCR can be used. The InterMusE archive has recently been moved into the Greenstone3 digital library, and as part of this process Google Vision OCR has been created for every scanned document, so that future versions will be able to use this directly (Bainbridge *et al.*, 2023). OcrMarkup uses word-level OCR and ignores larger phrase/line structures provided by Google or Tesseract, as text

---

[3] Note that 'annotation' here refers to the labelling of specific parts of an image or document. This typically includes some form of location in the document and an associated comment or note. In the case of TalkOver the location is a glowing spot and the comment/note the audio recording. In the case of OcrMarkup the location is a rectangle and the comment the name of the region.

```
// decide if current word runs on from last word
let bbx1 = bounding box of last word
let bbx2 = bounding box of current word
let lineHeight = maximum height of bbx1 and bbx2
if ( abs(bbx1.top-bbx2.top) > 0.6 * lineHeight ) {
    // if vertical position very different
    return false  // not a run
}
if ( bbx2.left - bbx1.right > 1.5 * lineHeight ) {
    // if horizontal gap too big
    return false  // not a run
}
return true       // yes a run
```

FIGURE 9. Pseudocode for finding word runs within OCR text

has to be re-threaded within selected regions. Instead a simple custom algorithm is used to detect co-linear text (see Figure 9), which gives near perfect results on all images tested to date.

OcrMarkup shares the same HTML framework as TalkOver for import/export of completed OcrMarkup annotation and pictorial browsing of past annotations. In both OcrMarkup and TalkOver, MD5 digests are calculated for all immutable media to make it easier to connect multiple annotations to the same underlying image.

5 Discussion – shared value

The two prototypes described here were designed and attuned to the specific contexts of the different communities. There are some common features, notably both are pseudo-geographic—they are associated with specific places, but the people do not live alongside one another. This means that community communication and coherence is through specific events and online means such as Facebook. The Troedrhewfuwch community, however, does have an identifiable, albeit uninhabited, patch of ground, whereas the music societies are intrinsically dispersed, and have always been so. While they both fit Ruth Finnegan's description of a group '*bonded by numerous ties, [who] know each other and have some consciousness of personal involvement in the locality of which they feel part*', in the (former BMS) music societies that would be truer of a committee member than of someone attending a concert for the first time and/or perhaps 'just passing through' (Finnegan (2007), p. 299).

The groups share an interest in community heritage preservation, but differ markedly in socio-economic terms, and more fundamentally in purpose. For the music societies their history is an essential part of their identity, but in the end it is secondary to their ongoing musical passion. For the Troedrhiwfuwch community, history and heritage are central to their activities and goals, but for most of them this is in a largely informal sense. Correspondingly the prototypes that arose from the two groups are very different. We can think of various stages of heritage archives: **collecting** primary and secondary material, **curating and organizing** this to enable future use, and finally **communicating** within and beyond the community. Both prototypes are focused on the first of these, collecting, but have a different tenor: TalkOver is focused on informal reminiscence, whilst OcrMarkup is more clearly archival in nature, reflecting the differing purposes and backgrounds of the communities and the co-production activities that gave rise to the bespoke designs.

The surprise, that perhaps should not have been a surprise, is what happened when each prototype was demonstrated to the other group.

When TalkOver was shown to the InterMusE academic team they immediately saw potential value and it was included in an upcoming meeting with music-society members. This was a very early version of TalkOver and it was hard to change the image used, so the demonstration was with a photograph of people from Troedrhiwfuwch (Figure 3). Despite the unfamiliar material, the music-society members also instantly saw potential applications, thinking particularly of long-standing members of the music society who could talk about old concert programmes or AGM minutes adding anecdotes, identifying people, and more. In addition, when an early version was presented at IAML,[4] TalkOver also generated considerable interest even though the prototype design was focused on non-professional users.

Following this, the OcrMarkup demo (again in early form) was presented to the Troedrhiwfuwch community. The document was the concert programme in Figure 8, so not a local document. This was partly due to the difficulty of changing the document, as at that stage the document was being parsed by hand through the web portal of Google Vision API. It was also less obvious how it would apply, as many important documents for the Troedrhiwfuwch community, such as census records or birth certificates, were hand-written. Perhaps because of this, there was no 'aha!' moment akin to that when TalkOver was demonstrated to the music societies.

A few months after this, however, the Troedrhiwfuwch community approached the research team to ask if the OcrMarkup application was available for use. A new and important document had been added to the community archive and they realized that this was the perfect tool to use for that.

In each case, the 'bespoke' tool custom-designed for the specific needs of a particular community turns out also to be of use to the other very different community. In addition, TalkOver is currently being considered for capturing community memories prompted by photographs in another project, the Willow Community Project, which is focusing on a legendary but now closed Cantonese restaurant-cum-disco in York (Hodgson, 2022). A full exploration of this use case will appear elsewhere, but suffice it to say for current purposes that TalkOver shows clear potential for generalization beyond these projects.

As noted, this perhaps should not have been surprising. Studies of 'single-person design', where an application has been targeted at a single individual, found that even the most personalised application was appreciated by others (Razak, 2008); indeed many successful web applications have arisen out of such situations, Wordle being perhaps the most recent example (Victor, 2022). Similarly, there are enough deep commonalities between apparently different communities that solutions targeted at one are of value to others.

This is very encouraging. There are many projects where universities have worked closely with community groups to create innovative prototypes for community heritage and communication (Beel *et al.*, 2017; Dix *et al.*, 2016; Taylor & Cheverst, 2009). However, if we really want to democratise digitization, to put tools for digital heritage into the hands of communities, we need to create *reusable* tools or, as Avram *et al.* (2020, p. 255) put it, a legacy

'beyond project-specific endeavours'. While at first this seems at odds with co-production, in fact our experience is that the creation of applications to help specific situations and the design of tools for general use can go hand in hand.

This is not to say that every tool designed for a specific community will be useful for all others, but for each targeted tool, there will be a number of other communities for which it is also a useful or even ideal solution. This has been explored at an individual level in *designing for peak experience* Dix (2010), which highlights the difference between 'good enough for all' designs, for universal use such as a word processor, compared to 'best for some' applications, such as game design. For these 'peak experience' applications, a viable and often the best development path is to optimise for an individual, and only when it is right for that person to attempt to generalise for a slightly larger group. We suggest that this is also a viable and maybe the preferable development path for communities also.

6 Current developments and future work

As noted, the InterMusE project digital archive is now in a custom installation of Greenstone3 digital library (Bainbridge *et al.*, 2023). Greenstone is a long-standing open-source digital-library platform produced by the New Zealand Digital Library Project at the University of Waikato. It has many stable installations worldwide and the latest version, Greenstone3, includes full IIIF support for document images (Bainbridge & Witten, 2020, Witten, 2009). As part of the custom ingest process, Google Vision OCR is performed on all scanned documents. An OpenAnnotation Server and Mirador Viewer (Sanderson *et al.*, 2015) have been installed alongside, making use of Greenstone's flexible extension mechanisms; together these allow each scanned-document image to be zoomed to high detail using IIIF and annotations to be added to regions of any scan.

For the next stages of work with the Troedrhiwfuwch community, there are plans to secure research funding, which will enable the collection and recording of stories and histories, using TalkOver and OcrMarkup, from the oldest surviving community members (currently in their 80s and 90s). It is recognised that these narratives are very fragile, and collection opportunities time-limited. Failure to collect and record these stories as soon as possible, before the last surviving older members of the community pass away, will mean they are lost forever. TalkOver will be used for some of this process, but we are also expecting some audio-only recordings. The oral-history team may like to add photo annotations to these audio recordings retrospectively; so TalkOver will be modified to allow this.

Also on the recording side, both for Troedrhiwfuwch village and also for the Willow project (see *Discussion* section), we wish to have TalkOver installed on the respective websites to allow members of the respective communities to record remotely. This will mean adding mechanisms to edit, upload and moderate content. Although we will be working closely with these projects, we will attempt to make this as self-maintaining as possible, probably by creating a WordPress plugin for TalkOver.

While designed for community-heritage purposes, these tools and many of the general lessons leading to them, also speak to emerging issues in traditional archiving projects. As noted, TalkOver generated interest amongst professional librarians and archivists at IAML; in addition we have since reflected on our own previous practice and realised further potential overlap. The InConcert project, which preceded InterMusE, was dealing with more standard scholarly archives of digitised material. This highlighted the need for more open and flexible approaches to the

---

[4] Panel, 'Opening up the digital archive: insights on openness in digitization and digital archiving from the InterMusE project', at the *International Association of Music Libraries (IAML) Conference*, Prague, July 2022. Panel Chair: Rachel Cowgill. In-person speakers: Rachel Cowgill, Charlotte Armstrong, and J. Stephen Downie. Video presentations by Alan Dix and Mike Twidale.

maintenance of material from sources of varying authority (Dix *et al.*, 2014b) an issue that cannot be ignored in community archives. Also, while the team was creating the online version of Brown and Stratton's 1897 *British Musical Biography* (Dix *et al.*, 2014a), a version of a tool such as OcrMarkup would have avoided much painstaking work.

7 Conclusion

There is always a temptation to try and develop a one-shot one-size-fits-all application, including when designing an archival database. This may be reasonable for large-scale institutions, where procedures can be formulated and staff can be trained in particular practices and formats. For local communities, however, we must design for the needs and peculiarities of each. We may not know what people want to say about digitised artefacts until we give people the opportunity to tell us—and that, in turn, depends on widening the range of those who get to tell us things.

We therefore need to think about accessibility down to the level of the individual user. Both TalkOver and OcrMarkup demonstrate a relationship between accessibility, enhanced metadata, and enriching the historical record. Using OCR on a digitised document makes it searchable, but also perceivable to someone using a screen-reader. A concert programme, however, is highly complex (as discussed earlier): off-the-shelf OCR might convert the column of instruments and performers' names in Figure 8 into a single unpunctuated paragraph, which a human reader would find confusing, and a screen-reader would simply recite (see *web materials demo1.mp3*). If someone were to correct the OCR for this concert programme, and then use the OcrMarkup tool to specify the content of the different areas, we would get something that is not only a lot more useful in terms of data, but that also generates a much more usable output for screen-readers (see *web materials demo2.mp3*). So we need to be thinking about how we can get the most out of the data these types of interventions will gather, for as many users as possible.

We also need to design for the unexpected. We may not know what we will find in the archive until we have finished digitizing—as was seen in the delayed realization of the potential of OcrMarkup by the Troedrhiwfuwch community. This means that as well as designing for particular needs right now, we also need to design for ease of revision (refactorability), to make it feasible and affordable to redesign to accommodate future needs and use scenarios. In general, we may not know how the database of digitised artefacts will be used in the future, or how and why it might get interconnected with myriad other databases with all kinds of different content. This underlies our own use of flexible semantics and annotation, but we are aware that we need also to find easy ways to modify these and/or connect them to external ontologies and authority files.

Looking at the two prototypes, while these are developed for different needs and purposes, they also share common features. Both are focused on annotation of images: one linking pictorial/photographic images to added audio commentary, and the other linking textual areas to named attributes. If the same programme were semantically annotated and also had TalkOver stories, we might want to be able to search annotations by name and then use this to index stories that point to the faces of the named people. In some ways this is rather like facets of an underlying semantic model. One could create a general 'do it all' application for media annotation, or indeed select one that already exists, but that would lose the specific qualities and simplicity that make each tool work. Our challenge is to find ways to

have multiple targeted applications that share sufficient common data representation to enable sharing and linking, yet are still flexible enough to make entirely new co-produced applications possible.

We are doing all of this in the context of community heritage and, more widely, historical archives. However, we are also aware that many of the issues we face in looking at this larger picture of connection, curation and annotation are shared in other domains, for example data analysis. We hope that by keeping focused on our own domain, we also create concepts and solutions that may be useful more widely—just as the communities we have described here found uses for the tools designed for each other.

More information on the projects and prototypes described here can be found at: https://www.alandix.com/academic/papers/IwC2024-community/.

References

Anderson, B. *Imagined Communities: Reflections on the Origins and Spread of Nationalism*. Verso, 1983.

Armstrong, C., Cowgill, R., Dix, A.,Bashford, C., Downie, J. S., Twidale, M., Reagan, M. and Ridgewell, R. (2021) Towards a foundation for collaborative digital archiving with local concert-giving organisations. In Arthur, C. (ed), *The 8th International Conference on Digital Libraries for Musicology*, pp. 41–49. ACM.

Armstrong, C., Cowgill, R., Dix, A., Bashford, C., Ridgewell, R., Reagan, M., Twidale, M. and Downie, J. S. (2023) Reframing ephemera: digitisation, community music-making, and archival value(s). In Giglitto, D., Ciolfi, L., Lockley, E., Kaldeli, E. (eds), *Digital Approaches to Inclusion and Participation in Cultural Heritage*, pp. 160–180, Routledge.

Arrandale, K. (2023) The International Musician 1922–1926. *Edward J. Dent: A Life of Words and Music,* chapter 10. Boydell & Brewer.

Avram, G., Ciolfi, L. and Maye, L. (2020) Creating tangible interactions with cultural heritage: lessons learned from a large scale, long term co-design project. *CoDesign*, **16**, 251–266.

Bainbridge, D., Cowgill, R., Perry, F., Downie, J. S., Dix, A. and Twidale, M. (2023) Collaborative musicology: Designing a digital library of musical events ephemera. In Thomas, M. E. (ed), *DLfM 2023:*

Proceedings of the 10th International Conference on Digital Libraries for Musicology, pp. 119–127. ACM.

Bainbridge, D. and Witten, I. H. (2020) A renewed look at Greenstone: lessons from the second decade. *IPSI Trans. Internet Res*, **16**, 3–11.

Beel, D., Wallace, C., Webster, G., Nguyen, H., Tait, E., Macleod, M. and Mellish, C. (2017) Cultural resilience: the production of rural community heritage, digital archives and the role of volunteers. *Journal of Rural Studies*, **54**, 459–468.

Brown, J., D. and Stratton, S. S. (1897) *British musical biography: a dictionary of musical artists, authors, and composers born in Britain and its colonies*. SS Stratton.

Cohen, A., P. (1985) *Symbolic Construction of Community*. Ellis Horwood / Tavistock Publications.

Cousins, M. E. (1935) A Historic Music Event. *The Music of Orient and Occident: Essays towards Mutual Understanding,* chapter XIII. B.G. Paul, Madras.

Cowgill, R. (2018) Relations with the enemy: Arthur Eaglefield Hull, musical internationalism and the politics of postwar recovery. Paper presented at 'A "Musical League of Nations"? Music Institutions and the Politics of Internationalism' conference. Institute of Musical Research. Senate House, London, 29–30 June 2018.

Cowgill, R. (2022) Musical internationalism and post-war cultural regeneration: Arthur Eaglefield Hull and the British Music Society of 1918. Paper presented at the 21st *Quinquennial Congress of the International Musicological Society*, Athens, 22–26 August 2022.

Cowgill, R. and Holman, P. (eds) (2007) *Music in the British Provinces, 1690–1914*, Ashgate.

Craig, G. and Mayo, M. (2011) *The Community Development Reader: History, Themes and Issues*. Policy Press.

Crow, G. and Mah, A. (2012) Conceptualisations and meanings of 'community': the theory and operationalization of a contested concept. Technical report, University of Southampton and University of Warwick.

Dix, A., Cowgill, R., Bashford, C., McVeigh, S. and Ridgewell, R. British Musical Biography online. (Online version of J. Brown and S. Stratton, 1897). *InConcert Project*. Accessed 14 Nov 2023. https://datatodata.com/in-concert/BMB/, 2014a–.

Dix, A. (2010) Human–computer interaction: A stable discipline, a nascent science, and the growth of the long tail. *Interacting with Computers*, **22**, 13–27.

Dix, A., Cowgill, R., Bashford, C., McVeigh, S. and Ridgewell, R. (2014b) Authority and judgement in the digital archive. *Proceedings of the 1st International Workshop on Digital Libraries for Musicology*, 1–8.

Dix, A., Cowgill, R., Bashford, C., McVeigh, S. and Ridgewell, R. (2019) Crowdsourcing and scholarly culture: understanding expertise in an age of popularism. In Khan, V.-J., Papangelis, K., Lykourentzou, I., Markopoulos, P. (eds), *Macrotask Crowdsourcing: Engaging the Crowds to Address Complex Problems*, pp. 189–214. Springer.

Dix, A., Jones, E. B., Neads, C.-A., Davies, V., Cowgill, R., Armstrong, C. B., Ridgewell, R., Twidale, M., Downie, S. R., Reagan, M. and Bashford, C. (2022) Tools and technology to support rich community heritage. In 35*th Int. BCS hum.-comput. interact. conf.*, pp. 1–10. British Computer Society.

Dix, A., Malizia, A., Turchi, T., Gill, S., Loudon, G., Morris, R., Chamberlain, A. and Bellucci, A. (2016) Rich digital collaborations in a small rural community. In Anslow, C., Campos, P., Jorge, J. (eds), *Collaboration Meets Interactive Spaces*, pp. 463–483. Springer.

Finnegan, R. (2007) *The Hidden Musicians: Music-Making in an English town* (2nd). Wesleyan University Press.

Giglitto, D., Lazem, S. and Preston, A. (2018) In the eye of the student: an intangible cultural heritage experience, with a human-computer interaction twist. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, pp. 1–12. ACM.

Giglitto, D., Lazem, S. and Preston, A. (2019) A participatory approach for digital documentation of Egyptian bedouins' intangible cultural heritage. *Interac. Des. Architect. J.*, 31–49.

Hodgson, E. *The Willow: Gone but not Forgotten*. Accessed 29/11/2023. https://www.streetlifeyork.uk/discover/willow-gone-not-forgotten, 2022.

Hoyle, V. (2022) *The Remaking of Archival Values*. Routledge.

Hutchinson, H., Mackay, W., Westerlund, B., Bederson, B. B., Druin, A., Plaisant, C., Beaudouin-Lafon, M., Conversy, S., Evans, H., Hansen, H., Roussel, N. and Eiderbäck, B. (2003) Technology probes: inspiring design for and with families. In *Proc. of the SIGCHI conf. on hum. factors in comput. syst. (CHI '03)*, pp. 17–24. ACM.

Kelly, B. L. (2023) Internationalism and musical exchange in post-World War I Europe, 1918–1923. In Fléchet, A., Guerpin, M., Gumplowicz, P., Kelly, B. L. (eds), *Music and Postwar Transitions in the 19th and 20th Centuries*, Berghahn.

Kirby, S. (2023a) The British (and International?) Music Society: Australian identity and musical internationalism in interwar Sydney. *Twentieth-Century Music (forthcoming)*.

Kirby, S. (2023b) 'To foster the spirit of international music': Louise Dyer and the British Music Society. In Murphy, K., Hill, J. (eds), *Pursuit of the New: Louise Dyer as Music Publisher and Collector*, pp. 35–54. Lyrebird Press.

Lambert, D. and Frisch, M. (2019) Digital curation through information cartography: a commentary on oral history in the digital age from a content management point of view. *The Oral History Review*, **40**, 135–153.

Lave, J. and Wenger, E. (1991) *Situated Learning: Legitimate Peripheral Participation*. Cambridge University Press.

Masters, G. (2021) *New-Music Internationalism: The ISCM Festival, 1922–1939*. PhD thesis. Music Department, King's College London.

Nutbeam, D. and Kickbusch, I. (1998) Health promotion glossary. *Health Promotion International*, **13**, 349–364.

Oral History Society. The history of oral history. Accessed: 2022-02-07. https://www.ohs.org.uk/about-2/, 2022.

Razak, F. H. A. (2008) Single Person Study: Methodological Issues. PhD thesis, Computing Department, Lancaster University.

Robertson, B. (2021) Lace version 0.5 new features (draft of 2019-12-21). Technical report, Mount Allison University.

Sanderson, R., Snydman, S., Winget, D., Albritton, B. and Cramer, T. (2015) Mirador: a cross-repository image comparison and annotation platform. *Proceedings of the 10th International Conference on Open Repositories (OR2015)*.

Sharpless, R. (2008) The history of oral history. In Charlton, T. L., Myers, L. E., Sharpless, R. (eds), *Thinking About Oral History. Theories and Applications*, pp. 7–32. Altamira Press.

Shelemay, K. K. (2011) Musical communities: rethinking the collective in music. *Journal of the American Musicological Society*, **64**, 349–390.

Social Care Institute for Excellence (2015) Co-production in social care: what it is and how to do it. Accessed 29/11/2023. https://tinyurl.com/5c9et5wz.

Studdert, D. and Walkerdine, V. (2016) Being in community: re-visioning sociology. *The Sociological Review*, **64**, 613–621.

Taylor, N. and Cheverst, K. (2009) Social interaction around a rural community photo display. *International Journal of Human-Computer Studies*, **67**, 1037–1047.

UNESCO (2003) Convention for the safeguarding of the intangible cultural heritage. Text available at: https://ich.unesco.org/en/convention.

Vayanou, M., Katifori, A., Chrysanthi, A. and Antoniou, A. (2020) Cultural heritage and social experiences in the times of covid 19. *AVI²CH@ AVI*.

Victor, D. (2022) Wordle is a love story. In *New York Times*, 3 January 2022.

Walkerdine, V. and Studdert, D. (2012) Concepts and meanings of community in the social sciences. Discussion Paper. AHRC Connected Communities.

Wenger, E. (1999) *Communities of Practice: Learning, Meaning, and Identity*. Cambridge University Press.

Willmott, P. (1986) Social networks, informal care and public policy. Research Report 655. Policy Studies Institute.

Wilson, J. M. (1882) *The Gazetteer of Scotland*. W. & A.K. Johnston, (Digital version 2020. https://alandix.com/gzs1882/).

Witten, I. H. (2009) The development and usage of the Greenstone digital library software. *Bulletin of the Association for Information Science and Technology*, **35**, 31–38.

Zembrzycki, S. (2013) Bringing stories to life: using new media to disseminate and critically engage with oral history interviews. *Oral History*, **41**, 98–107.