Emergent archives and crowdsourced narratives: two development stories from the Queensland State Library

Seth Ellis

Griffith University, South Brisbane QLD 4101, Australia
seth.ellis@griffith.edu.au

Abstract. This paper suggests and connects two ideas: First, that digitising collection objects creates new objects in the form of the digital records, which can then form the basis of emerging archives in patterns of public use and interaction. These emerging archives then become a powerful vector for community participation and decentralised authorship of historical and cultural narratives. As such, they deserve to be collected, archived, and made public themselves, which may require new archival strategies.

Second, I suggest that digitising non-material objects such as audio-visual materials allows us to examine new ways of describing and cataloguing historical material, by using narrative as metadata. This use of narrative description as part of the essential cataloguing of objects is also of use in organising and understanding community contributions to catalogues and descriptions. Narrative metadata is one of the particular affordances of digitised archives and collections, and can be used not only to strengthen community engagement, but to generate new, archivable historical material in the form of public narrative contributions.

The confluence of these two ideas is apparent in collections like that of the State Library of Queensland in Brisbane, Australia, which houses community object-archives in the form of home movies, photos, and other materials collected from or donated by the public. I explore these ideas through two recent projects at the State Library: the Corley Explorer, and my own work as 2019 Mittelheuser scholar-in-residence, exploring sound as historical material.

Keywords: digital affordances, community engagement, narrative, metadata, sound, emerging archives.

1   Community participation in information structures: the Corley Explorer

This paper examines the affordances of digitized archives—specifically, that such an archive creates other archive-worthy materials in the form of community interaction. Fostering these interactions in a way that allows their contributions to fully and productively enrich the archive itself, we must re-examine the way we use archives and make them accessible. In following this idea I will examine two recent projects at the State Library of Queensland in Brisbane: first, the Corley Explorer, a public interface for a collection of 61,000 photos of Queensland houses from the 1960s and 1970s;

second, my own research as 2019 Mittelheuser scholar-in-residence at the State Library, working on the idea of sound as historical material in the Library's existing collections.

Archived media such as photographs can function as themselves archives of memory, as Van Alphen suggests:

> The photograph is not only an archival record, it is also an archive in itself. The photographic image is a spatial configuration of one moment, a configuration that consists of a great number of details. All these details or elements are stored in the image in order to be ordered or classified, leading to one reading or another. Conversely, such use of photographs further illuminates the archive and makes its principles more widely visible and thus common. Acknowledging this double, mutual relationship between photography and the archive, one can say that, in the words of Philip Monk, photographs are lodged and lodging at the same time. [1]

A photography archive is an archive of archives, but in most cases the secondary archive—photo-as-repository—remains invisible, uncatalogued, and undescribed. This secondary archive is necessarily subjective, consisting of personal and social histories, observations, and interpretation. In order to treat the photo-as-repository as an archive, we need to treat these narratives as a type of metadata.

This "doubling" relationship of archived photographs — and, I will argue, of other mediated, audiovisual representations in collections and archives — can be recognised not only in academic writing, but in museological practice, through the implementation of community-sourced narrative information, as is being done in various forms around the globe. The *Waisda?* project of the The Netherlands Institute for Sound and Vision, for instance, used crowdsourced annotations to augment the institute's own efforts to make their collection objects more accessible to the public.[2] The bulk of these efforts are to support existing cataloguing structures, and in some cases offer new tags and other details of searchable metadata. I suggest that going further, and collecting these strategies into a single conceptual framework of "emergent archives" — the archives of personal or social histories embedded in a collected object such as a photograph, and the interactions of the public with it — is a useful way to understand and strengthen the particular affordances of the digital archive.

In 1995, the State Library of Queensland acquired the Corley Archive, a collection of 61,000 photos. Frank and Eunice Corley had spent more than a decade in the 1960s and 1970s taking photos of houses around suburban Southeastern Queensland, and selling them to homeowners. Two-thirds of their photos were sold; the collection acquired by the Library constitutes the remainder.

For twenty years this enormous mass of photo-documentation remained in Library storage, rarely accessed and mostly unseen. In 2015 the Library began a massive effort to digitise all the photos—with what metadata existed, which was only a few handwritten notes from the Corleys. The Library worked with the Annerley Stephens History Group and the University of Queensland's *Architecture Theory Criticism History Research Centre* to identify what houses they could, and what stories could be gleaned

from them. The result was, first, the *Home* exhibition in the Library, which ran from December 2018 to July 2019, showcasing the photos and their interpretations in various ways; and second, the Corley Explorer, an online tool for crowdsourcing further identification of houses in the photos. [3]

The Corley Explorer was developed by Mitchell Whitelaw and Geoff Hinchcliffe at the Data Design Lab, ANU School of Art & Design. Their original purview from the State Library was to address community accessibility of the Corley Archive, and also to create a structure whereby community members could assist in filling out the metadata for specific photos in the collection. The resulting online interface went public upon the opening of the Home exhibition on 7 December 2018. It mapped as many photos as possible to their physical location; it also allowed logged-in users, members of the public, to add information in several ways: by placing on the map, by adding tags, and by contributing stories.

Following the State Library's previous experience with community volunteer efforts, the Library expected to get a relative handful of history enthusiasts contributing their time in particular to tagging photos. To this end, they suggested tags that had mostly to do with the architectural nature of the houses: "pitched roof," "single storey," "weatherboard walls," and so forth. This strategy is of especial significance in Queensland, whose "Queenslander" cottages make up one of the very few regional vernacular architectural styles in Australia. This effort has been very successful thus far, surpassing the Library's expectations; over the first weekend after the Explorer's launch, more than 3000 people visited and contributed to the Explorer's information.[4] Over time, as is usual in the case of open-participation strategies, the contributions in tagging and placing have evolved into considerable efforts by a relative handful of "superusers" who do much of the work, and occasional efforts by other members of the public.[5]
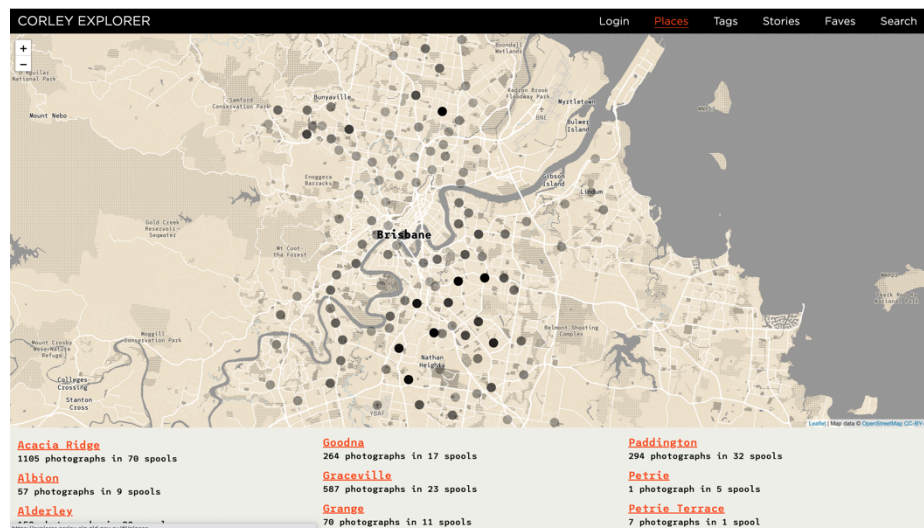

Fig 1: Corley Explorer, showing photos whose subject have currently been mapped

Less expected was the enthusiasm in the public for contributing and sharing stories about the houses represented in the Archive. As of this writing, 2041 photos have had stories attached by members of the public—a tiny fraction of the total number of photos, but a considerable collection in itself.[6] Many of these stories are purely informational and indeed architectural, for instance:

> **68 Pozieres Road, Tarragindi**
> *Story added on 26 September 2020 by Ken A*
> From memory this was the original house for the area bounded by Pozieres Rd, Messines Ridge Rd and Queensthorpe St. Possibly a soldier settlement block, possibly for a Mr Gibson.[7]

In other cases, the histories take the form of memories:

> **61 Power Street, Norman Park**
> *Story added on 24 September 2020 by Roberta M*
> The man who owned this house always grew the most beautiful flowers along the front street boundary. He would always be there weeding and watering his flowers and would cut a few stems for me to put in a vase in my room as a child. He took great pride in his property.[8]

> **21 Lochel Street, Mount Lofty**
> *Story added on 29 August 2020 by Gary L*
> I used to visit here as a young child, the house belonged to my Grandparents. the [sic] were owners from before WW2 to early sixties. I can remember Grandma starting the wood stove on cold mornings. I'm never asked about its history. I know my great grandparents lived in a House named Maritana in Stuart St.[9]

In yet other instances, multiple histories add layers of both information and affect:

> **9 Tatong Street, Indooroopilly**
> *Story added on 19 December 2018 by Jennifer C*
> Prior to the Western Freeway construction this house was located in Julie Street. Julie Street now ends on the other side of the Freeway.

> *Story added on 24 September 2020 by Amanda R*
> I grew up in this house from the moment my parents bought me home from hospital as a new born until I was 9 years old. It was a great little area and as young kids we used to explore the foothills of Mt Cootha. The original address was 49 Julie St. My Dad and Mum were very proud of their first house and kept the gardens beautiful and did many renovations. Sadly when the Western Freeway was built, our street was cut in 2 and the whole dynamic of the neighbourhood changed. A new development

was built next door and suddenly the quiet little enclave was no more. But what fun we had as kids - especially watching the freeway get built![10]
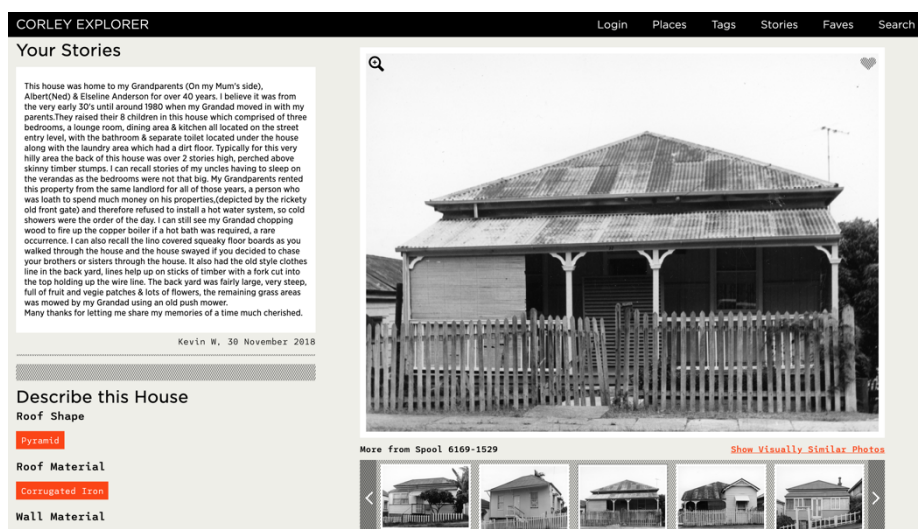
Fig 2: A single photo's entry in the Corley Explorer, showing a few tags and considerable reminiscence. The address is 32 Plunkett St., across the road from the author's house.

Within the first year of community contributions, it was clear to Whitelaw and Hinchcliffe, as well as to the librarians at the State Library, that the contributed stories themselves constituted a considerable archive, and that they thus deserved to be preserved and *to be accessible* in the manner of any other archive. While the stories are of course online, and are thus in theory accessible to anyone with an Internet connection, in practice the stories are not deeply searchable or structured. A Corley Explorer 2.0 is currently under development, that will more thoroughly store and structure the emergent archive as searchable metadata.[11]

A question of considerable importance in this is the relationship between community-contributed information, including the emergent archive of personal recollections, and the "official" archive of Library acquisitions. Should the database of user-contributed material, and the database of Library-owned objects and their metadata, be kept entirely separate? Does community-contributed information show up in the Library's official search engine (OneSearch)? Part of the Library's strategy in outsourcing tagging is to bring the various Corley images into alignment with existing tag structures within their collections; architectural features such as "corrugated iron roofs" are existing subjects within their collections. However there is currently no mechanism for incorporating narrative metadata into the Library catalogue and its descriptions, and unlike some of the Library's other online exhibitions, photos displayed in the Corley Explorer do not link back to their OneSearch catalog entries. The Corley Explorer and the official catalogue are designed to be entirely separate user experiences, though the

librarians use the Explorer to add metadata to the catalogue, within the structure the catalogue already contains. Nonetheless the Explorer has two clear, successful uses: to bolster the cataloguing activities of the library, and to create and maintain an ongoing conversation of storytelling amongst community members. This dual modality informed my later thinking as Mittelheuser scholar-in-residence, as I describe below.

Fig 3: The OneSearch entry for one of the images in the Corley Archive, showing none of the emergent information from the Corley Explorer.

2   Object description in audio-visual materials: sound in the State Library's collections

In 2019-20 I was the Mittelheuser scholar-in-residence at the State Library of Queensland, part of the Library's Queensland Memory Awards. My project was "sound as historical material: developing a new way of cataloguing, describing and accessing sound in the archive." Over the course of this project, now concluded, I developed and suggested to the Library strategies for representing sound materials in their archived collection—both digitised, and yet-to-be-digitised—that would make those sounds better represented, more accessible, and more useable by the public. This means re-thinking to a degree how we treat sound as a material in the first place, and how we catalogue it. By the end of this residency it became apparent that the technical, formal modes of cataloguing I found fruitful would benefit in practical application—indeed, might be made possible—through an application of the community-based tagging and narrative contextualisation we saw in the Corley Explorer. This involves more storytelling than simple tagging, and thus creates more cultural information that could, in potentia, itself be archived.

This project arose out of my interest in sound as an affective trigger of past experience, and in using sound as a means to create more evocative narratives of historical place. Sound is both a visceral connection to the past, and the most unknowable thing about the past [12]. On the one hand, sounds can cause immediate, phenomenological re-living of past events in the present. Far from being intangible, sound is one of the few historical materials that touches the participant; it enters your ear. On the other hand, sounds last only for moments, and the majority of sounds made by humans have never been recorded. When we address sounds beyond performance and language, it opens up the possibility of critically and imaginatively addressing the spaces of human life, and the way that we occupy them.

This paper does not suggest a novel method of opening library catalogues to community input, or of radically restructuring these foundational back-stage informational structures to admit of more and more types of metadata on a case-by-case basis. I suggest however that one of the positive aspects of digital collection is community participation; that this participation can take the form of storytelling; and that it will be difficult to fully embrace the power of this distributed, crowdsourced storytelling unless we evolve pragmatic strategies to treat narrative as metadata. Digital archives are foundationally different from physical archives, and the objects they create — digital entries — offer different affordances not only for surface-level community participation, but for deep cataloguing, description, and accessibility. In this second part of the paper I suggest that narrative as metadata is also a useful model for re-cataloguing audio-visual materials, by treating the digitised mediations of those objects as durational materials that can be annotated and "transcribed" in a model borrowed from the digitisation of oral histories.

Ala Rekrut has written extensively on "material literacy" in archival studies—that is, the way that sensitivity to the material aspects of records within an archive can also reveal information about the past:

> The physical object is already a conceptual, as well as a physical, "data" object. It is a physical site where a variety of kinds of information reside the most obvious being the written text and images on the surface of the records – but it also bears information about itself as a record-object, about the text and images on its surface, and about the culture which produced, cared for, and used it…
>
> Changes to the records and to their meaning is ongoing, for as David Lowenthal notes: "artifacts are simultaneously past and present; their historical connotations coincide with their modern roles, commingling and sometimes confusing them ... . The tangible past is in continual flux, altering, ageing, renewing, and always interacting with the present. [13]

More lately, Burns has extended the concept of material literacy to suggest that it ought also to encompass digitised archives; the digital version of an object, she suggests, has its own nature, history, and presence.

> As a marker, the digital image itself is an additional form of metadata for the object, as it captures and preserves not only the content but also the current state of its material existence at a particular moment in time. The digital image solidifies the object in time. Because the original and its digital counterpart age differently, the surrogate essentially provides two versions of the same singular object. [14]

These writers are dealing largely with the digitisation of physical objects, such as photographs. Sound materials in collections go one step further; the original "sound object," the sound-as-made, necessarily lasts for only a few seconds; if it exists in collection at all, it is by definition *already mediated* by recording technology. Unlike other, "tangible" artifactual objects, sound can only be preserved as a copy of itself.

The requirements of an "object" that exists only as mediation can be deduced from the nature of sound materials as described in many collections' catalogues, including that of the State Library of Queensland. In the case of many entries, the media object is identified as containing sound, but the sound itself is not identified. In other cases, the presence of sound is hidden behind one-word entries like "film;" there is no further information whether the film has sound attached. The most usual result is, it is possible to search for sound, but not possible to search for *particular* sounds.

There are a number of currently available AI-driven platforms for parsing sound files and making them searchable. However these engines—for instance, Google Cloud—are built largely on the assumption that the sound files contain either language (as in oral history) or musical performance [15]. In other words, these are performed sounds. What of non-performative sounds, what Murray Schafer famously called *keynote sounds* that might evoke a "soundscape" in the sensory imagination of the listener? [16] This can be an important affective record of the past, but contemporary technology does not recognize and cannot contextualise these sounds. How could an algorithm recognize, for instance, the sound of a mangle, once common on laundry day?

If sounds in the broadest sense—musical, vocal, and other—cannot be digitized and made searchable in the same way as text, how can we make the contents of audio-visual materials visible to the catalogue, and thereby accessible to public search and use? In the Mittelheuser scholarship I suggested treating the audio-visual object as an archive in itself, in the manner Van Alpern identifies, and in the way the Corley Archive successfully experimented with. Because audio-visual materials are time-based, however, we need a different strategy for exploring that secondary, object-based archive. The *Waisda?* project successfully trialed the use of crowdsourced annotations in adding crucial descriptive metadata, not only augmenting the limited time available to curators but also introducing useful and descriptive tags beyond the conventional tags that made up the curators' toolkit.[17] That project attached metadata to the video file as a whole, as an archived object. For the State Library I suggested an annotative structure adapted from platforms like Soundcloud, that allow for commentary and description within a given audiovisual material. That is, the specific sound becomes the object, not the audiovisual file as a whole.
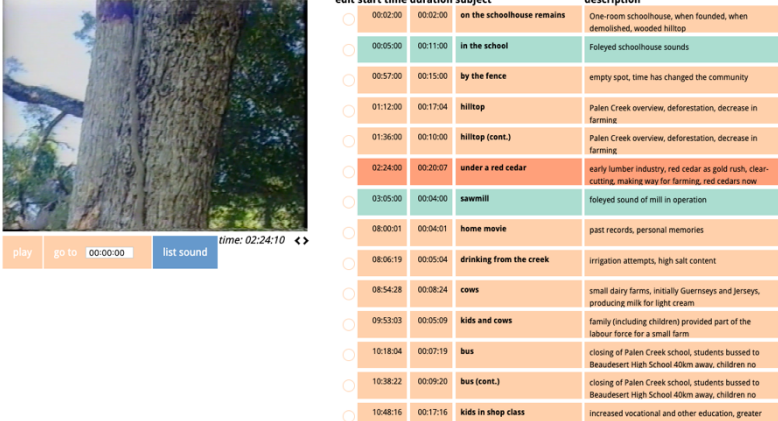
In beginning the Mittelheuser project, I adapted techniques from the catalogue descriptions of oral histories, specifically at Trove, to suggest that individual audio-visual objects could be "transcribed" in the manner of oral histories, no matter what the sounds. As an initial test case I used a documentary in the John Oxley Collection, *Palen Creek: A Changing Community* by John Moloney, and "catalogued" each non-vocal sound in its run time.

Fig 4: An oral history entry in Trove, with transcription and summary [18]

29993 Gary Moloney short films 1994-1995
Palen Creek: A Changing Community

sounds

← back to media collection

edit	start time	duration	subject	description
○	00:02:00	00:02:00	on the schoolhouse remains	One-room schoolhouse, when founded, when demolished, wooded hilltop
○	00:05:00	00:11:00	in the school	Foleyed schoolhouse sounds
○	00:57:00	00:15:00	by the fence	empty spot, time has changed the community
○	01:12:00	00:17:04	hilltop	Palen Creek overview, deforestation, decrease in farming
○	01:36:00	00:10:00	hilltop (cont.)	Palen Creek overview, deforestation, decrease in farming
○	02:24:00	00:20:07	under a red cedar	early lumber industry, red cedar as gold rush, clear-cutting, making way for farming, red cedars now
○	03:05:00	00:04:00	sawmill	foleyed sound of mill in operation
○	08:00:01	00:04:01	home movie	past records, personal memories
○	08:06:19	00:05:04	drinking from the creek	irrigation attempts, high salt content
○	08:54:28	00:08:24	cows	small dairy farms, initially Guernseys and Jerseys, producing milk for light cream
○	09:53:03	00:05:09	kids and cows	family (including children) provided part of the labour force for a small farm
○	10:18:04	00:07:19	bus	closing of Palen Creek school, students bussed to Beaudesert High School 40km away, children no
○	10:38:22	00:09:20	bus (cont.)	closing of Palen Creek school, students bussed to Beaudesert High School 40km away, children no
○	10:48:16	00:17:16	kids in shop class	increased vocational and other education, greater

play | go to 00:00:00 | list sound *time: 02:24:10* ‹ ›

Fig 5: A web app I created to identify and describe non-vocal sounds in audio-visual materials.

This granular approach to sound materials proved useful in identifying the contents of long time-based media, and in making those contents accessible both through simple search and through auto-generating tags from commonly used identifiers. It became apparent early in the development process that this mode of cataloguing was only deeply useful, however, with the presence not only of the subject matter, but of the description. Historical sound becomes an evocative measure of the past when we are able to contextualise it.[19] In the case of Moloney's documentary, this was easy to achieve, since Moloney helpfully narrated the meaning of his sounds and visuals himself. Thus the sound of a shop classroom becomes the sound of small dairy farms in the area losing their children to good-paying industrial jobs; the "sound" of a lone red cedar in a field evokes the massive forests of red cedar that fell prey to the logging industry, now defunct itself.

The narrative of historical context was an important component of making sound affective in its evocation of the past. But in other materials the sound was less easily described by those who were not there at the time. Even the visuals of a home movie, for instance, are often mysterious except at the shallowest level. We see the dirt yard of a country station, but without prior experience we don't know what's just off camera, how near its neighbours are, whether they more often ride horses or in cars, et cetera. In fact, the sense of what's happening outside the visual frame is what we often use sound for, to create a fuller sense of place. But without narrative context we don't know how to interpret the sound either; and in some cases there is no sound, or rather, the sound is the unnatural silence of soundless film media.[20]

By the end of this residency it became apparent that the best way to approach the contextualisation and description of sound materials would be to crowdsource it, in the way proven effective by the Corley Explorer. Community members might be able not only to identify, but to contextualise and narrate, individual scenes and sounds in a way

that is not possible by myself or the Library staff. Thus the newest version of my cataloguing strategies and platform, developed in parallel with the Corley Explorer 2.0, suggests a framework in which information can be added by, and exchanged amongst, members of the public. This public contribution might be suggested by specific inquiries from library staff or from others: thus the construction of a category of contribution as "questions," highlighting the degree to which investigation of the past is an inquisitive and collaborative act. The aim is to use gaps in the catalogue—indeed, in the manner catalogues are constructed—as spaces that can be filled with public authorship and storytelling, within a structured archival setting.

Fig 6: A new version of the web app allowing for additions from the public.

There are considerable historical tensions in the GLAM sector between the two imprimaturs of archival collections: archiving and accessibility.[21] Digitising collections is of course one way to improve accessibility, though only to a mediated version of the original collected object. Much of the materiality of the original object—the paper and ink of diary entries, the celluloid of home movies—remains hidden from view. What digitised archives do afford, however, is the ability not only to create new public interactions but to create new knowledge emerging from those interactions. Community participation can consist not only of letting community members in the door—whether the door is physical or virtual—but of soliciting new archival material in modes that will, increasingly, need to be structured according to community's interest and ability to participate. Using emergent community narratives as archivable material, tying this new archive to an existing collection archive on a deep structural level, is one strategy that can provide exciting new possibilities in how institutions create and store knowledge. The question remains: what is the relation of this new archive of digital ephemera to the existing archive in the Library's collection, and its metadata? At the moment the Library is choosing to keep these two structures separate: one as a repository, one as a living conversation. The next step might be to use the affordances of digital structures, not to incorporate one of these archives into the other, but to create easily navigable paths between them, so that community members can move freely back and forth between the two in a single afternoon's browsing. Whether to store and safeguard the emergent archive against time remains a question for the future.

References

1. Van Alphen, E.: *Staging the Archive: Art and Photography in the Age of New Media.* University of Chicago Press, Chicago (2014).
2. Oomen, J. et al.: "Waisda?: making videos findable through crowdsourced annotations." (2014). In: Ridge M. (Ed.), *Crowdsourcing Our Cultural Heritage,* pp. 161-185. Taylor & Francis Group, London (2014).
3. Corley Explorer Homepage, https://explorer.corley.slq.qld.gov.au/, last accessed 2020/9/27.
4. Gavin Bannerman, director of Queensland Memory, in conversation 2018/12/13.
5. Hinchcliffe, G., Whitelaw, M.: From data to meaning. Making Meaning symposium, State Library of Queensland, Brisbane (2020/2/6).
6. Corley Explorer Stories page, https://explorer.corley.slq.qld.gov.au/#!/stories, last accessed 2020/9/27
7. Corley Explorer Stories individual entry, https://explorer.corley.slq.qld.gov.au/#!/photograph/6169-2951-0016, last accessed 2020/9/27
8. Corley Explorer Stories individual entry, https://explorer.corley.slq.qld.gov.au/#!/photograph/6169-2567-0004, last accessed 2020/9/27
9. Corley Explorer Stories individual entry, https://explorer.corley.slq.qld.gov.au/#!/photograph/6169-0589-0004, last accessed 2020/9/27
10. Corley Explorer Stories individual entry, https://explorer.corley.slq.qld.gov.au/#!/photograph/6169-1425-0017, last accessed 2020/9/27
11. Hinchcliffe, G., Whitelaw, M.: From data to meaning. Making Meaning symposium, State Library of Queensland, Brisbane (2020/2/6).
12. Riello, G: The material culture of walking: spaces of methodologies in the long eighteenth century. In: T. a. R. Hamling, Catherine (Ed.), *Everyday Objects: medieval and early modern material culture and its meanings*, pp. 41-56. Ashgate Pub ,Farnham, Surrey, England (2010).
13. Rekrut, A.: Material Literacy: Reading Records as Material Culture. *Archivaria 60* 11-37 (2005)
14. Burns, J. E.: The Aura of Materiality: Digital Surrogacy and the Preservation of Photographic Archives. *Art Documentation: Journal of the Art Libraries Society of North America 36*(1), 1-8 (2017).
15. Orio N.: Searching and Classifying Affinities in a Web Music Collection. In: Agosti M., Bertini M., Ferilli S., Marinai S., Orio N. (eds) Digital Libraries and Multimedia Archives. IRCDL 2016. Communications in Computer and Information Science, vol 701 (2017).
16. Schafer, M.: *The Soundscape: Our Sonic Environment and the Tuning of the World.* Inner Traditions/Bear & Co. (1993)
17. Oomen, J. et al. (2014)
18. John Rigby interview at Trove, https://nla.gov.au/nla.obj-215479961/listen, last accessed 2020/9/27.
19. Riello (2010)
20. Altman, R.: *Silent Film Sound*. New York: Columbia University Press.
21. OCLC Research: *Making Archival and Special Collections More Accessible*. Dublin, Ohio (2015).