



The Dorothy Howard Collection: revealing the structures of folklore archives in museums

Mike Jones^a, Kate Darian-Smith^a, Deborah Tout-Smith^b and Gavan McCarthy^c

^aSchool of Historical and Philosophical Studies, The University of Melbourne, Victoria, Australia; ^bHome and Community Life, Museums Victoria, Victoria, Australia; ^ceScholarship Research Centre, The University of Melbourne, Victoria, Australia

ABSTRACT

In 2015–16 researchers from the University of Melbourne and Museums Victoria undertook a collaborative project which sought to visualise archival data from the museum as a means to investigate the structure and context of the Dorothy Howard Collection. This article introduces Dorothy Howard's work, which is part of the internationally significant Australian Children's Folklore Collection, and looks at the project's processes and outcomes, including the initial visualisations produced. In doing so, the authors highlight the data-intensive nature of such work, suggest the potential of visualisations to reveal collection structures, and outline possibilities for future projects and collaborations.

KEYWORDS

Archives; museums; folklore; collection management; documentation; networks

You put down on a piece of paper as many dots as you please. You must not do this just any way or you will be wrong. You must go straight up and down and straight across so that the dots will be in squares. Then you take turns joining lines between the dots.

(Frank Syaranamual, 10 years old, describing a game called 'Dots', 1954)1

Large collections in archives and museums may hold many thousands of items,² none of which exist in isolation. If each item is taken as a dot, or node, they are joined by complex networks of connections – to other items as well as to things such as people, organisations, concepts and places – with these relationships accumulating through time. Some items are explicitly joined, their content referring to other items or entities, while others are linked to context within or beyond their collections in ways which help users understand their meaning.

However, these networks are often not readily visible to users – whether they be internal staff of the organisation holding the collection, or external users exploring content. Archival description tends to emphasise hierarchical links between items, aggregates (series, collection) and an often-constrained set of provenance entities;³ and museum description tends toward individual item-level descriptions, loosely coupled to donor information or to like items via common characteristics and classification structures. Furthermore, even where

documented relationships exist, the structures they create can be difficult to see when items are viewed as discrete records in databases or online.

This article introduces the Australian Children's Folklore Collection (ACFC), and outlines the context and details of a proof-of-concept project which sought to explore the possibilities of producing visualisations based on relational networks. Supported by the McCoy Seed Fund, the project involved Museums Victoria staff working with archivists and historians from the University of Melbourne on the Dorothy Howard Collection, which sits within the ACFC. A key challenge of the project was the extraction and curation of data from museum and archival documentation (collection metadata) in a systematic manner that would enable digital humanities scholars and other researchers to use existing visualisation tools to study the collection. It was not the purpose of the project to undertake analyses of the resulting network graphs to any significant degree, or to transcribe, extract and use data contained in the records themselves, but to prepare the way for others.

The results provide valuable insights into the structure of the Dorothy Howard Collection, and indicate the amount of work required to prepare existing catalogue data for visualisation. The project also helps reveal something of the ways in which museum and archival systems interoperate (and some of their limitations in this area), shows the potential of network graphs to support the navigation and interpretation of collections, and suggests future possibilities for working with documentary collections held by museums.

The Australian Children's Folklore Collection

The Australian Children's Folklore Collection (ACFC) contains a range of material culture concerned with the history of childhood and children's play. 'Children's folklore' is defined by Museums Victoria as the folklore both of children and about children. It recognises that such lore - from games to verbal culture - is created and transmitted by both children and adults.⁶ The ACFC is Australia's pre-eminent national collection of children's folklore material, and one of the largest and most significant collections of its kind anywhere in the world. In recognition of its significance, the collection was placed on the UNESCO Australian Memory of the World Register in 2004, sitting alongside such key documents of national and international historical significance as the Endeavour journal of James Cook, the Mabo Case Manuscripts, and the landmark constitutional documents underpinning the Australian Constitution.8

The ACFC was established in the 1970s, when children's folklore specialists June Factor and Gwenda Davey began amassing and analysing a range of objects and documents pertaining to children's play drawn from their own research, and from that of their students. This collection was later acquired by Museums Victoria, and has been supplemented by items collected in other contexts. Today the ACFC consists of over 10,000 file cards, recordings, objects, and other written and visual materials relating to children and their play worlds. These include subcollections such as interviews with adults and children in Aboriginal communities across Victoria about play9; materials collected in the 1980s relating to the 'multicultural' playground; and play artefacts from around the world, as well as Australia. The ACFC also holds the archives of several projects researching play that were conducted from the 1970s onwards. Among these are written, aural and visual records from the recent major national study Childhood, Tradition and Change, which documented schoolyard play activities across the nation and situated these in historical perspective. ¹⁰

The Dorothy Howard Collection is one of the largest components of the ACFC, and contains rich documentation of the lives and play of Australian children in the middle of the twentieth century. At its core are the fieldwork notes and research findings of the noted children's folklorist Dorothy Howard, who visited Australia from the US for 10 months during 1954–55. Funded as a Fulbright scholar, Howard was initially based at the University of Melbourne, where she found there was little interest in both women scholars and the study of children's play culture. There was, however, unprecedented interest during the decades after World War II in the importance of children to national development as future citizens and workers. The rising birth-rate as a result of the 'baby boom' and the launch of the large-scale immigration scheme to Australia had, by the mid 1950s, resulted in a dramatic increase in the number of children within the population. To address a desperate shortage of facilities, all Australian states embarked on large building programs during the 1950s and 1960s to provide schools, libraries, and other community and sporting facilities aimed at young people. 11

During Howard's time in Australia she conducted intensive fieldwork observation of children's games and play activities in primary school playgrounds. She travelled widely, visiting every state (with the exception of the remote Northern Territory), documenting children's play in a variety of settings, from one-room schoolhouses in rural areas to overcrowded inner urban schools. She invited schools she could not visit to post her written descriptions of play, and in many cases teachers compiled students' work to pass on to her. She also interviewed and received correspondence from adults about their childhood memories, in some cases creating a record of play that stretched back to the nineteenth century. Howard's papers and fieldwork notes are also held in the ACFC, with details of particular activities – such as running and skipping games, hopscotch and counting-out rhymes – transcribed in her handwriting onto small index cards.

Howard was interested in evidence of regional variances in play, whether verbal chants and rhymes or the conduct of games. Her analysis also remarked on the influence of British and American popular culture on Australian children. At a time when television had not yet been introduced in Australia (it came in 1956, just following Howard's visit), she found such 'foreign' influences were generally gleaned from films and magazines aimed at boys and girls, including comic books, which were imported at the estimated rate of 60 million per year and were seen to have a negative effect on Australian youth.¹² Howard also documented the rapid change in play objects brought about by new technologies, such as the shift from the use of real sheep knucklebones to ones made from plastic. On her return to the United States she wrote up her findings on Australian children's play practices and on customs that continued from a century earlier, such as keeping autograph albums. Originally published in academic journals, these analyses were republished in 2005 as a collected volume accompanied by contextual and explanatory research essays.¹³ Together these pieces complement the raw data in the ACFC collection.

The Dorothy Howard Collection is a valuable resource for folklorists and historians. It is also interesting as a largely documentary collection held by a major museum, one of only a handful of collected archives held and managed along with Museums Victoria's artefacts. The contents are an insight into the world of children in the mid twentieth century, often incorporating their voices and words as opposed to the adult and institutional voice. From the perspective of collection description and management, the records of this collection sit at the intersection of museums and archives, the items functioning simultaneously as artefacts and documents, many of which need to be understood as part of an interconnected network of schools and correspondents with Howard at its centre. It is therefore a collection which is ideally suited to the approach outlined in this article.

Managing and documenting archives in Museums Victoria

Since the late 1970s, museums have shown an increasing interest in the archival collections held by their institutions. The emergence of the so-called 'museum archives movement' in the United States led to many long-standing institutions receiving funding and assigning resources to help them gain physical and intellectual control over distributed records. 14 From the early 1980s there have also been a number of publications which focus explicitly on the development of archival and recordkeeping programs in museums, 15 these produced alongside initiatives from professional associations, including the Society of American Archivists' Museum Archives Roundtable, and the Standing Conference on Archives in Museums in the United Kingdom.¹⁶

But from the inception of the museum archives movement, the focus was on internal separation from other collections. The single-page draft guidelines produced at the first Conference on Museum Archives, held at the Smithsonian's Belmont Conference Center in December 1979, specified the following: 'The museum archives is a repository in which are preserved museum records ... When practical the archives should be a separate department ... The archives should be located in a separate and secure area of the museum.¹⁷

Many museums now separate their archives physically and document them using different systems and standards.¹⁸ Museums Victoria is no exception. Following the establishment of a formal archives and records management program in 1998, the organisation began physically centralising many records, subsequently purchasing a licence for records management software in 1999-2000 to help document and manage these collections. 19 Over the same period the museum invested heavily in the development of KE EMu (Knowledge Engineering's Electronic Museum, now known simply as EMu),²⁰ the collections management software still used today to manage and describe the State Heritage Collection.

The distinction between the 'museum' and the 'archives' is not always clear, particularly in the case of documentary material, and is based more on provenance than format or descriptive requirements. Field books are an example. When written by a museum staff member, field books usually become part of the institutional archives, while those produced by external naturalists and anthropologists are kept as part of the State Heritage Collection.²¹ Similarly, as Dorothy Howard was not a museum staff member, her index cards and other field notes are not considered part of Museums Victoria's archives, nor are they documented as such.

When the ACFC was formally established in 1979, the National Museum of Victoria (which merged with the Science Museum of Victoria in 1983 to form the Museum of Victoria, now Museums Victoria) had no archivists and no formal archival program. Although the Belmont conference included 'acquired (non-institutional) records'22 in the remit of museum archives, Museums Victoria has focused its archival work primarily on organisational records; 'external' collections like Dorothy Howard's remained where they were.

As a result, the Dorothy Howard material is documented in EMu as a collection of 'museum' items (as are all 'archival' items in the State Heritage Collection at Museums Victoria), with some archival elements incorporated into the data structure. (While EMu has archives tabs and fields, these are currently not used by the museum.) Thirteen series have been identified as part of the ACFC, including Series 3, the Dorothy Howard Collection (HT 8468). Within this are a number of 'subseries' (or child records in EMu terms), including Manuscript Files, 1954–55 (HT 8468.1) and Personal Correspondence (HT 8468.9), with individual items registered within those subseries (making them grand-child records).

However, though hierarchically linked as 'parent-child' records, series and subseries structures are not presented as part of online collection search, and information is repeated at item level. For example, many items contain the same (or similar) information in their summary description, as shown in this text attached to a piece of correspondence:

One of a collection of letters describing a children's games written to children's Folklorist Dorothy Howard between 1954 and 1955. Dr Howard came to Australia in 1954-55 as an American Fulbright scholar to study Australian children's folklore. She travelled across Australia for 10 months collecting children's playground rhymes, games, play artefacts, etc. This letter, together with the other original fieldwork collected by Dr Howard during this period, is preserved in the Dorothy Howard Collection manuscript files, part of the Australian Children's Folklore Collection (ACFC), Archive Series 3. The ACFC is an extensive collection documenting children's folklore and related research.

Furthermore, this text is part of a summary created at page level (not attached to files or other recordkeeping units), isolating individual sheets of paper from their broader context and – from an archivist's perspective – providing little understanding of collection structure or original order. The description is therefore a mixture of parent-child records and itemlevel description within EMu, and for online users the Dorothy Howard Collection appears as a large number of discrete item-level entries.

While museum collection management systems present challenges for archival description, they can also be used to capture a wealth of fielded data in ways often not supported by archival tools. The record for the transcribed letter from Frank Syaranamual (a quote from which opens this article) includes collection and acquisition information; author, addressee, creator and collector details; named organisations; physical description; and fields for classification, category, discipline and keywords.

EMu also supports the documentation of relationships between content. The descriptions of 'Charades' and 'Dots' from August 1954, featured on a single typed page, were initially collected by Dorothy Howard as handwritten letters from Syaranamual (aged 10) and Hazel Vost (aged 11). Both are in the collection, ²³ and have been linked to the typed page along with articles about the Australian Children's Folklore Collection and Dorothy Howard.²⁴

Each page can therefore be associated with a rich set of data, and the archival aggregates, item-level descriptions, authors and named organisations create numerous cross-references and relationships. However, such detail is rarely captured as part of the daily work of cataloguers, and where such data exists it can be difficult to discern when viewed either as an archival hierarchy or as discrete museum records. In the case of the Dorothy Howard Collection, the authors and their colleagues recognised an opportunity to not only digitise the collection and record this information, but also to explore the potential of network visualisation tools as a means for representing this relational data.

Enhancing and visualising the Dorothy Howard Collection

In late 2014 a \$20,000 grant was received from the McCoy Seed Fund to extend documentation, access and insight into the ACFC, with a particular focus on the Dorothy Howard Collection. The Fund is a joint initiative between the University of Melbourne and Museums Victoria to foster innovative and high-impact collaborative research projects, leveraging the strengths of the two partners to build scholarship based on Museums Victoria's collections.

The project was guided by an inter-agency committee under the leadership of coordinating investigator Gavan McCarthy, Director of the eScholarship Research Centre (ESRC). Its interconnected aims were to:



- · demonstrate how innovative documentation systems, visualisation tools and other technologies can enhance collections held by institutions;
- make collections more discoverable and understandable to researchers and wider users;
- demonstrate the potential of visualisations as a means by which to explore, and potentially access, collection data;
- provide the foundations for future researchers to navigate and analyse the complex relationships between creators, themes and places in the Dorothy Howard Collection;
- contribute to the ongoing development of the visualisation and database technologies involved.

Intensive curatorial work was undertaken by Georgia Knight, who worked closely with Collection Information Management Systems project officer Christina Giatsios to identify data standards and maximise the use of bulk loading to rationalise data inputs. The close teamwork between curatorial and information systems staff was key to the success of the project, ensuring high data quality and streamlined processes throughout.

The McCoy grant work on the Dorothy Howard Collection included registration to individual document level, description, contextual research, images, keywords and cross-referenced Subjects and Parties (people, organisations). Among the project's achievements was the uploading of thousands of orphan images to EMu and the physical rehousing of the entire Dorothy Howard Collection of 4600 documents to museum best practice. At the conclusion of the project, a total of 7732 EMu catalogue and multimedia records had been created, 711 EMu catalogue records had been enhanced and 539 new catalogue records were released to Collections Online (http://collections.museumvictoria.com.au/), providing digital access to large numbers of individual documents in ACFC for the first time. In addition, 481 EMu party records (for people and organisations) had been created, and a total of 12,507 EMu records had been linked.

The research team knew that, once the data was captured, they wanted to visualise the content. For some time the ESRC has explored tools designed for this purpose, including the open-source Java desktop application Cytoscape (http://www.cytoscape.org/) and the ESRC's own web services-based Contextual Network Explorer (ConneX: https://connex. esrc.unimelb.edu.au/#/>), which uses a Pyramid web application to construct graphs and a client-side interface written in AngularJS to visualise the results. While the former can produce visualisations from basic Excel spreadsheets and comma-separated values files, ConneX has been developed for archival data – specifically a form of Extensible Markup Language (XML) known as Encoded Archival Context - Corporate Bodies, Persons and Families (EAC-CPF). 25 Though there have been some experiments with EAC-CPF in museums, including at the American Museum of Natural History,²⁶ there are no established processes for converting EMu data to EAC-CPF and associated records.

Therefore, the project team decided to export data from EMu for import into the ESRC's Online Heritage Resource Manager (OHRM), a standards-based contextual information management system based in Microsoft Access, which could then be used to create EAC-CPF files for visualisation in ConneX as well as the more basic data required for Cytoscape. ESRC and Museums Victoria staff spent time working out a suitable cross-map between EMu and OHRM fields, and Nancy Ladas (Manager of Collection Information Systems at the museum) arranged exports of data from EMu as Excel files structured in a way which would facilitate import into the OHRM with only a small amount of manual manipulation. Over the course of the project 1019 records (539 catalogue records and 480 party records) were exported to the OHRM, along with 2874 relationships.

Network visualisation

Understanding the complexity of the world through graphs – nodes connected by links, or relationships - has a history dating back to the work of Leonhard Euler in the early eighteenth century.²⁷ In more recent years analyses of the Internet, social network theory and the idea of 'small worlds'28 have all contributed to the use of networks to map and visualise the way people, groups, communities and other elements interconnect. Digital humanities scholars and practitioners have also explored the potential of network graphs and social network analysis while bringing humanistic understanding of the constructed nature of knowledge to their use and application, challenging notions of objectivity and working to reveal the biases inherent in algorithmically derived representations of data.²⁹ Drawing on this prior scholarship, network graphs are viewed by the authors as another way of seeing, but one which is not necessarily more truthful simply because the method starts with the mathematical manipulation of data in place of a human encounter with text.

Alongside these social and cultural investigations has been the gradual emergence of network-based views of cultural heritage collections. In the archival community Gavan McCarthy and Joanne Evans, Tom Lynch, and Daniel Pitti et al. have explored social and contextual networks as they relate to archives and archival description, including means for extracting and exposing these structures.³⁰ As archivist and academic Joanna Sassoon writes, 'archival thinking' is a way of working where 'the constituent parts of a cultural and physical landscape are part of a broader system. The broader system is a network of related entities interacting together with their non-living (physical, geographical and administrative) environment.'31 Museum and archaeological theorists Antonio Cordella, Fiona Cameron and Helena Robinson, Sarah Byrne et al. and Ian Hodder (among others) also explore the networks that exist between things in the world and how these relate to the structure and meaning of collections.³²

In this project the ESRC used both ConneX and Cytoscape to develop comparable network visualisations. A primary aim of these visualisations was to enable the collection to be seen as a whole, overcoming some of the limitations of catalogue interfaces, print guides and narrative descriptions. In the case of ConneX the same force-directed algorithm is used for all the dataset produced by the tool, establishing consistency and predictability of visual analysis. As ConneX only accepts EAC XML, a consistent data format standard, and uses a single algorithm, it is possible for the user to compare the outputs of different collections and develop their skills in reading these graphs. By completely restricting user enhancement of the network outputs (except through filters and colouring of nodes), ConneX limits the tendency of users to drift towards 'infographics' (i.e. message-directed visualisations) rather than focusing on understanding the relationship between the data and the algorithm.

When using Cytoscape the ESRC primarily uses the same approach, applying the yFiles Organic Layout Style:

The organic layout style is based on the force-directed layout paradigm. When calculating a layout, the nodes are considered to be physical objects with mutually repulsive forces, like, e.g., protons or electrons. The connections between nodes also follow the physical analogy and are considered to be springs attached to the pair of nodes. These springs produce repulsive or attractive forces between their end points if they are too short or too long. The layout algorithm simulates these physical forces and rearranges the positions of the nodes in such a way that the sum of the forces emitted by the nodes and the edges reaches a (local) minimum.³³

Though, as with all algorithmically generated representations of data, this results in a limited perspective - just one way of seeing out of many possibilities - force-directed layouts are useful starting points as they are designed to clearly represent complex networks, including knowledge networks.34

Aside from a basic visualisation of the Dorothy Howard Collection, key series within that collection, and Howard herself, the next step was to use ConneX to examine the documented collection down to item level. There is great value in seeing an entire collection in a single image, something impossible to achieve in any meaningful way using text, lists of content, indexes or other representational methods. Used in this way, network graphs are a means by which to view and understand the structure of the collection itself, as documented by collection metadata. Other types of visualisations - maps, timelines or images derived from textual analysis – were beyond the scope of this project. The records have not been transcribed or put through optical character recognition processes; and, even if this additional level of data had been available, these other visualisation methods would focus on the content of the records themselves at the expense of the structure of the collection.

The print format of this article limits the utility of showing what are in some cases large and complex visualisations. As a consequence, only a small number of figures have been included in the following section.

Figure 1 utilises the relationship data from the EAC-CPF records to visualise the Dorothy Howard Collection (indicated), and the five groups of records that make up Series 3 (circled):

- Manuscript Files, 1954–55;
- Personal Correspondence, 1954–55;
- Publications by Dorothy Howard, 1937-69;
- Poetry and Creative Writing by Dorothy Mills (Howard) and her pupils;
- Interviews and Lectures, 1954–55.

The largest of these 'subseries', circled in the top left corner of Figure 1, is Manuscript Files, 1954–55, which contains 456 documents, shown as the large mass of nodes between it and Dorothy Howard, whose node is – not surprisingly – at the centre of the network.

Even with comparatively basic visualisations such as these where structure and connectedness are evident, some humanities scholars may find them alien and impenetrable. Extracting meaning can require interpretation, which in turn requires training and familiarity before people become confident and fluent in their use.³⁵ For example, though centrality, node size and clustering can all be understood as expressions of the impact and significance of particular nodes within the broader network, sometimes the resulting image may initially appear counterintuitive. In Figure 1, the Dorothy Howard Collection (indicated) is much smaller than Howard's large central node. While the collection's centrality suggests (quite rightly) that it has connections to all 'sides' of the graph, it is dwarfed by Howard's node because the collection is only connected directly to a small number of entities - the series it contains, and Howard - and not to all the items within those series. These are second-order relationships and therefore not considered when calculating node size.

Interpretative problems become even greater when the full network is revealed, as shown in Figure 2 (produced using Cytoscape). Here the nodes are a uniform size with centrality a key indicator.

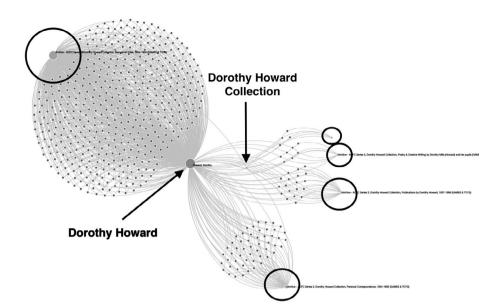


Figure 1. Force-directed visualisation of the Dorothy Howard Collection, created using ConneX. eScholarship Research Centre, 2016.

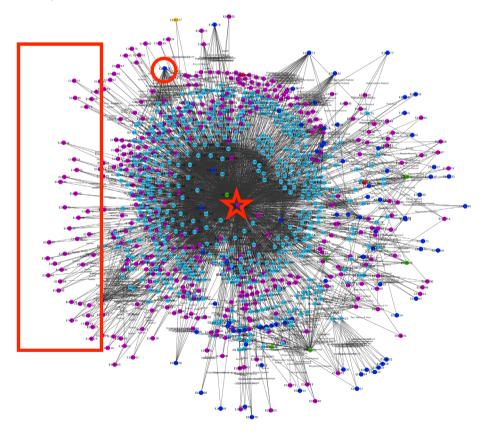


Figure 2. Force-directed visualisation of the Dorothy Howard Collection, created using Cytoscape. eScholarship Research Centre, 2017.

Once again, the Dorothy Howard node is at the centre of the network (indicated by the star), surrounded by an inner ring of nodes mostly representing documents and a small number of people. When viewed in colour, people appear as purple, documents as light blue, organisations as dark blue and collection elements in green. The majority of 'people' nodes are on the outer edges of the network, such as those in the rectangle included as part of Figure 2. Howard had a large number of correspondents, but only a few with whom she exchanged more than one or two letters. Closer in are the schools with which she worked, for example Errol Street Primary School (circled), which is related to 42 documents in the collection. Clicking on the school reveals these documents (shown in yellow in Figure 3). However, Howard is not connected directly to either her correspondents or to organisations. Their connection is via documents - the archival items - sent by people or containing elements of children's folklore observed at particular schools.

Although this type of network graph (colloquially known as a 'hairball') is dense, it is still possible to discern structure, particularly when explored in colour on a screen, allowing the user to zoom in and out of each node to see labels and connections in detail. Such network graphs of 'small worlds' also tend to be relatively consistent in having a small

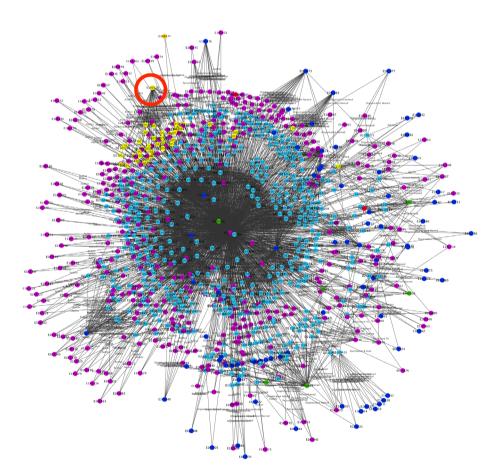


Figure 3. Force-directed visualisation of the Dorothy Howard Collection, with Errol Street Primary School and related nodes highlighted, created using Cytoscape. eScholarship Research Centre, 2017.

Table 1. Number of relationships to each entity, for entities with five or more relationships.

Identifier	Name	R/ships
E057181	Dorothy Howard	1103
E000002	Series 3 – Manuscript Files, 1954–55	457
E102676	Double View Government Primary School	106
E102675	East Camberwell Girls' Secondary School (College)	102
E102660	Collier Government Primary School	78
E099224	Unidentified	55
E000005	Series 3 – Personal Correspondence, 1954–95	54
E102678	Errol Street Primary School	42
E103320	Clayfield College	33
E102673	Carlisle Government Primary School	25
E103036	Ceduna Higher Primary School	24
E102683	East Fremantle Government Primary School	21
E000004	Series 3 – Publications by Dorothy Howard, 1937–69	20
E102696	George S. Browne	15
E103276	Cherry Gardens School	13
E000003	Series 3 – Poetry & Creative Writing by Dorothy Howard and her pupils	8
E093909	Charles P. Mountford	7
E102687	Brian Sutton-Smith	7
E103029	A.G. Paull	7
E103640	Coromandel Valley State School	7
E104068	Glenys Henderson	7
E000001	Dorothy Howard Collection	6
E103053	T.L. Robertson	6
E103165	Christine Brown	6
E000006	Series 3 – Interviews & Lectures, 1954–55	5
E103344	Patricia Watson	5
E103904	Brenda Hope	5
E104013	Pat Wingrove	5

number of highly connected nodes and a large number of nodes with only a few connections. Table 1 shows the top 28 nodes in terms of connectedness. Dorothy Howard has over 1100 relationships to other nodes in the network, but by the third node connectedness has reduced by an order of magnitude. Not shown in this table are the 453 items which have fewer than five relationships, including 328 with only one relationship.

Further work is required to see if this network, and others like it, have generalisable distributions, and to determine the amount of variation between networks. It is also important to note the distinction here between collection-centred networks and networks designed to capture general social relationships. In the context of children's folklore, this may include connections formed by site visits, the collection of photographs and artefacts, and the circulation of correspondence and related material. While many of the people involved may have other connections - whether by professional association or shared familial and friendship groups – these are not documented.

Folklore archives

When considering the broader implications of this work, collections of folklore are a potentially rich source of material for researchers who want to navigate and analyse the complex relationships that exist between documents and their context.

Folklore research includes a long tradition of creating and working with archives, one which folklorist Janet C Gilmore suggests is 'almost as old as the field of Western European archiving itself.36 While Gilmore's claim looks like over-reach, when Dorothy Howard travelled to Australia (and more than 25 years before the museum archives movement) the management of these collections had become a prominent topic in the field. Symposium II at the prominent Four Symposia on Folklore (1953) was dedicated to archiving folklore, leading the chair of the first session, Maud Karpeles, to conclude: 'Personally I found it extraordinarily cheering and stimulating to learn that the archiving of folklore is a very special and thrilling activity and is not the dull detection of dead things.³⁷

Subsequent to the symposia, the awareness that there was a need for more consistent practice led to a US journal, The Folklore and Folk Music Archivist, 38 with a 10-year print run (1958-68). By the 1980s, just as many American museums were starting to develop archival programs for the first time, folklorist Joseph C Hickerson noted, 'as fieldwork is a folklorist's basic tool, so is archiving.³⁹

Furthermore, since the late 1960s, structure and context have played an important role in the understanding of folkloric 'items' (here using that term in reference to individual tales, songs and sculptures, not to their documentation or material collection components). As folklorist Dan Ben-Amos wrote in his influential piece 'Toward a Definition of Folklore in Context':

folklore is not thought of as existing without or apart from a structured group. It is not a phenomenon sui generis. No matter how defined, its existence depends on its social context, which may be either a geographic, linguistic, ethnic, or occupational grouping.⁴⁰

In locating individual records as part of a visualised network of related entities (including schools, which fit with the idea of the 'small group' as the context of folklore, championed by Ben-Amos), the approach taken also retains an interest in the record as text. That is, the items are placed in their sociocultural context while retaining their status as documents with relationships to other documents in an archival sense. 41 Combining text and context in this way is a small step toward spanning the division between the two.⁴²

Use of network graphs for folklore collections is not new. As Timothy R Tangherlini points out in his introduction to a recent special issue of the Journal of American Folklore on 'computational folkloristics', there is a proliferation of such images used to analyse many different types of interconnected data, contributing to the development of the 'Folklore Macroscope'. Tangherlini sees this as 'a study environment that allows researchers to explore folklore data at varying levels of engagement. 43 But such work depends on well-structured data, something which can rarely be taken as a given. As noted in the introduction, the purpose of our investigation was not to undertake substantial analysis, but to develop the conditions under which such analyses become possible.

Project outcomes and future possibilities

As a result of the McCoy-funded project outlined in this article, Museums Victoria now has a highly developed process for documenting and physically managing archival collections. The level of resources required for intensive archival documentation in EMu has been determined; and an initial (though manual) process for mapping and exporting data from EMu into the OHRM has been established. Importantly, EMu has been identified as an appropriate environment for storing and managing detailed data on collections. While the OHRM provides an EAC-CPF export for ConneX, the system was not used to add data to that already existing in EMu. Future projects will likely explore means for visualising data directly out of EMu; and some initial discussion has taken place regarding the availability

of ConneX for Museums Victoria use, though it is likely adapting the tool would require additional funding.44

However, as this project has highlighted, whatever tools are in place, the data required to produce effective network visualisations is rarely available as part of collection documentation. Much of this study was spent on enhancing existing data, creating new content and capturing relationships, all with the aim of visualising the results (as well as making content available online via the museum's more established digital channels). While some items and collections include relational information, it is often either sparse and inconsistent, or stored in narrative text fields, making it difficult to access and parse.

The initial outcomes of the project highlight the value of undertaking such work, at least for high-use collections. For example, the ability to see the collection as a whole, and within that structure to quickly identify schools where a large amount of content was collected, or where very little activity occurred, highlights structural elements of the data which are difficult to see clearly via other means. Similarly, the potential to view Dorothy Howard and her correspondents (and, by extension, correspondence) as relational networks centred on Howard herself provides a new perspective on her work in 1954–55, revealing the structures of a self-contained network constructed through the gathering of children's folklore texts, images and artefacts.

Among the projects aims, noted earlier, was to contribute to the ongoing development of the technologies involved. Using ConneX to analyse the collection data highlighted known constraints of that tool. Most significantly, the reliance on EAC-CPF records without reference to Encoded Archival Description (EAD) records - in this case, documents - meant that only seven nodes were seen as 'connected' with 472 nodes seen as 'unconnected', which in the version of ConneX available meant that they were not expressed. These latter are, in fact, well connected into the network, but only via items. For example, Carlisle Government Primary School is linked to 21 documents and four articles, but though these items are themselves related to other people and organisations - including Dorothy Howard - and a more direct relationship can therefore be inferred, ConneX treats this school's node as 'unconnected'. That is, it is not directly connected to a person or another organisation. Owing to this, the Cytoscape visualisations, though not as interactive, provided additional insight. The project therefore provided evidence that the further development of ConneX as a tool for examining item-rich (rather than context-rich) datasets is both feasible and worthwhile.

Further research is required in other areas too, if the potential of these early findings is to be realised. For example, the structures revealed by network visualisations could be used as part of checking data quality and identifying outliers, or as a means for revealing and inferring new data relationships, some of which could then be captured in EMu. Other possibilities include: the use of network graphics as part of collection navigation and exploration, as seen with the Provisualizer on the recently updated Public Record Office Victoria website (https://www.prov.vic.gov.au/explore-collection-soil); as an integrated part of displays in exhibition spaces; or in conjunction with research platforms like those provided by the Humanities Networked Infrastructure Virtual Laboratory (https://huni. net.au/-/search>). The significant data enrichment work completed as part of this project, and the additional material now available online (and, as a consequence, via the Collections application programming interface, or API, available at https://collections.museumvictoria. com.au/developers>) makes the Howard data available to other researchers who can analyse and visualise aspects of the collection using their chosen tools.

Crowdsourcing remains a rich opportunity for future documentation of the ACFC, although the plan to crowdsource data enhancement during the project could not be undertaken in view of the high level of resources used to create quality EMu data and the lack of suitable technology through Museums Victoria's Collections Online. Fortunately, intensive curatorial work during the project produced more and higher-quality data than anticipated.

Further opportunities for work after the end of this project include the linking of the data to the Childhood, Tradition and Change Australian Research Council project undertaken in 2007-10, during which scholars from the University of Melbourne, Deakin University and Curtin University, in association with the National Library of Australia and Museums Victoria, collaborated to produce the first national account of continuity and change in Australian children's playlore since Dorothy Howard's work in the 1950s. The data from that project is also held in the OHRM, and analysis of changes across place and time using the visualisation tool presents a rich opportunity for research.

The project clearly underlines the potential research value of establishing relationships between otherwise isolated datasets. Other Australian studies of childhood which remain isolated include research by Peter Lindsay and Denise Palmer in Queensland in 1975-76, covering 21 Brisbane primary schools and observing nearly 5000 children. Potential international connections include Dorothy Howard records held at the Albin O Kuhn Library at the University of Maryland, Baltimore County, Texas and the University Libraries at the University of North Texas, as well as international collections such as the British Library, which holds a significant collection of children's folklore within the Opie Collection of Children's Games and Songs and the Children's Playground Games and Songs in the New Media Age collection.

Finally, as noted above, there are extensive folklore archives around the world, many now as old as, or older than, the Dorothy Howard Collection. Documentary records and artefacts are combined, with folkloric practice heavily reliant on understanding and tracing the interrelations of texts, contexts, people, groups, events and locations through time. The possibilities for combining new approaches to collection documentation with network visualisation in support of such research are substantial.

Conclusions

The Australian Children's Folklore Collection, and within it the Dorothy Howard Collection, have been recognised as internationally significant. In their combination of archival records, artefacts and contextual information, they are also a rich resource through which to explore the requirements of expanded collection documentation and network visualisation.

The McCoy Seed Fund proof-of-concept project was a first attempt to look at the requirements and possibilities of such work. The resulting network graphs were useful starting points, highlighting structures and networks which are difficult to see via other means, but requiring more analysis and work before they can be used in other ways. They are therefore, as cultural heritage technologist Mia Ridge has put it: 'an outcome of an exploratory process', but 'not necessarily the best way to present the final product'45 (whatever that product might be). Perhaps most significantly, the project demonstrated the large amount of data enhancement and capture required to support such an approach, and confirmed some of the known limitations of the systems and tools used. These were valuable insights which, through the combined expertise of museum staff, archivists and historians, revealed new ways of looking at collections, archives and folklore material, and laid important foundations for future collaborations.



Endnotes

- 'Document Hazel Vost, Addressed to Dorothy Howard, Descriptions of Acting Game "Charades" & Paper & Pencil Game "Dots", Aug 1954', Museums Victoria Collections, August 1954, available at http://collections.museumvictoria.com.au/items/2105273, accessed 28 February 2017.
- 2. Throughout this article 'item' is used to refer to a material component of a collection, whether it be an archival record or museum artefact (or both simultaneously). Except where specified, the term should not be taken in the sense employed by folklorists, for whom an item is 'a particular instance of a folk expression – a rendering of a particular song, folktale, or narrative motif, for example. See JC Gilmore, 'Filling "An Immense Brain with Very Little in the Brain" for "Perpetual Memory": Folklore Archiving New and Old, Journal of Folklore Research, vol. 52, no. 1, 2015, p. 124, doi:10.2979/jfolkrese.52.1.99.
- 3. Despite the challenges of Chris Hurley and others to the archival notion of provenance, the majority of archival description continues to focus on individual creators or collectors at the expense of more complex structures. See Chris Hurley, 'Problems with Provenance', Archives and Manuscripts, vol. 23, no. 2, November 1995, p. 234.
- 4. The McCoy Project, 'McCoy Seed Fund', The University of Melbourne, 2016, available at http://research.unimelb.edu.au/work-with-us/funding/internal/mccoy-seed-fund, accessed 28 February 2017.
- 5. Museum Victoria was renamed Museums Victoria on 1 November 2016. To avoid confusion the name of the organisation at the time of writing - Museums Victoria - has been used
- 6. This definition is employed elsewhere, for instance in the Opie collection of Children's Literature at the Weston Library (Bodleian Libraries, University of Oxford).
- Moya McFadzean, 'Australian Children's Folklore Collection', Museums Victoria Collections, 2004, available at http://collections.museumvictoria.com.au/articles/24, accessed 28 February 2017.
- 8. National Committee of Australia, Memory of the World, 2016, available at http://www.amw. org.au>, accessed 24 August 2016.
- Kate Darian-Smith, 'Oral Histories of Childhood and Playlore: The Aboriginal Children's Play Project, Museum Victoria, Aboriginal History, vol. 32, 2008, pp. 146-50.
- 10. Kate Darian-Smith and Nikki Henningham, 'Final Report of the Childhood, Tradition and Change Project', 2011, available at http://ctac.esrc.unimelb.edu.au/objects/project-pubs/ FinalReport.pdf>, accessed 24 August 2016; Kate Darian-Smith, 'Australian Children's Play in Historical Perspective: Continuity and Change on the School Playground', International Journal of Play, vol. 1, no. 3, 2012, pp. 264-78; Gwenda Davey, Kate Darian-Smith and Carla Pascoe, 'Playlore as Cultural Heritage: Traditions and Change in Australian Children's Play', in Kate Darian-Smith and Carla Pascoe (eds), Children, Childhood and Cultural Heritage, Routledge, London and New York, 2013, pp. 41-54; Kate Darian-Smith, 'The Heritage of Australian Children's Play and Oral Traditions', Oral Tradition, vol. 28, no. 2, 2014, pp. 233–42.
- 11. See Kate Darian-Smith, 'Children, Families and the Nation in 1950s Australia', in Kate Darian-Smith and June Factor (eds), Child's Play: Dorothy Howard and Children's Folklore in 1950s Australia, Museum Victoria Publishing, Melbourne, 2005, pp. 19-40; see also Hannah Lewi and David Nichols (eds), Community: Building Modern Australia, UNSW Press, Sydney, 2010.
- 12. Darian-Smith, 'Children, Families and the Nation', p. 28.
- 13. Darian-Smith and Factor.
- 14. Ann Marie Przybyla, 'The Museum Archives Movement', in Deborah Wythe (ed.), Museum Archives: An Introduction, Society of American Archivists, Chicago, 2004, pp. 3-8; Bruce Smith, 'Archives in Museums', Archives and Manuscripts, vol. 23, no. 1, May 1995, pp. 38–47; Charlotte Brunskill, 'The History of Record Keeping in the UK Museum and Gallery Sector', in Charlotte Brunskill and Sarah R Demb (eds), *Records Management for Museums and Galleries*: An Introduction, Chandos Information Professional Series, Chandos, Oxford, 2012, pp. 1–34.
- 15. William A. Deiss, Museum Archives: An Introduction, Society of American Archivists, Chicago, 1984, available at http://hdl.handle.net/2027/uc1.32106005529877, accessed 4 May 2017; Wythe; Brunskill and Demb.



- 16. Przybyla, p. 5; Brunskill and Demb.
- 17. 'Conference on Museum Archives', Archives of American Art Journal, vol. 19, no. 4, 1979, p. 25.
- 18. See Michael Jones, 'Artefacts and Archives: Considering Cross-collection Contextual Information Networks in Museums', in Nancy Proctor and Rich Cherry (eds), Museums & the Web: Selected Papers and Proceedings from Two International Conferences, Museums and the Web, Silver Spring, MD, 2016, pp. 123–35, available at http://mwa2015.museumsandtheweb. com/paper/artefacts-and-archives-considering-cross-collection-knowledge-networks-inmuseums/>, accessed 28 February 2017; Michael Jones, 'From Personal to Public: Field Books, Museums, and the Opening of the Archives', Archives and Records, online publication, January 2017, 1-16, doi:10.1080/23257962.2016.1269645.
- 19. The museum awarded a contract for records management software in 1999-2000, believed to be for TRIM Captura, which later became HP TRIM following the product's acquisition by Hewlett-Packard. The original installation is noted (without a product being named) in the museum's annual report. See Museums Board of Victoria, '1999-2000 Annual Report', 2000, p. 17, available at https://museumvictoria.com.au/pages/2877/1999-2000/MoV text3. pdf>, accessed 4 May 2017.
- 20. EMu is built on the Texpress database engine, which 'supports object-oriented database structures utilising text attributes, multi-valued fields, nested tables and object references'. See 'Engine', EMu website, available at https://emu.kesoftware.com/support/training-courses/ emu-training/98-support/emu-documentation/database/about-texpress/423-engine>, accessed 4 May 2017.
- 21. Jones, 'From Personal to Public', p. 4.
- 22. 'Conference on Museum Archives', p. 25; see also Smith, p. 39.
- 23. 'Document Frank Syaranamual, Addressed to Dorothy Howard, Description of Paper & Pencil Game "Dots", 11 Aug 1954', Museums Victoria Collections, 11 August 1954, available at http://collections.museumvictoria.com.au/items/2105263, accessed 28 February 2017; Hazel Vost, 'Document - Hazel Vost, Addressed to Dorothy Howard, Description of Acting Game "Charades" & a Transcription of a Counting-out Rhyme, 25 Aug 1954, Museums Victoria Collections, 25 August 1954, available at http://collections.museumvictoria.com. au/items/2105265>, accessed 28 February 2017.
- 24. McFadzean; Richard Gillespie, 'Dr Dorothy Gray Mills Howard (1902-1996)', Museums Victoria Collections, 2011, available at http://collections.museumvictoria.com.au/ articles/10599>, accessed 28 February 2017.
- 25. Society of American Archivists and Staatsbibliothek zu Berlin, 'Development of the EAC-CPF Standard, Encoded Archival Context Corporate Bodies, Persons and Families, available at <http://eac.staatsbibliothek-berlin.de/about/development.html>, accessed 23 May 2015; see also Richard V Szary, 'Encoded Archival Context (EAC) and Archival Description: Rationale and Background', Journal of Archival Organization, vol. 3, nos. 2–3, January 2006, pp. 217–27, doi:10.1300/J201v03n02_16.
- 26. Iris Lee, 'EAC-CPF at the AMNH, Part 1 | Hidden Collections', 18 March 2014, available at http://images.library.amnh.org/hiddencollections/2014/03/eac-cpf-at-the-amnh-part-1/; Iris Lee, 'EAC-CPF at the AMNH, Part 2 | Hidden Collections', 16 May 2014, available at http:// images.library.amnh.org/hiddencollections/2014/05/eac-cpf-at-the-amnh-part-2/>, accessed 28 February 2017.
- 27. Albert-Laszló Barabási, Linked: How Everything is Connected to Everything Else and What it Means for Business, Science, and Everyday Life, Plume, New York, 2003, pp. 9-13.
- 28. See Duncan J Watts, Small Worlds: The Dynamics of Networks between Order and Randomness, Princeton Studies in Complexity, Princeton University Press, Princeton, NJ, 1999.
- 29. See, for example, Johanna Drucker, 'Humanities Approaches to Graphical Display', Digital Humanities Quarterly, vol. 5, no. 1, March 2011; and Miriam Posner, 'What's Next: The Radical, Unrealized Potential of Digital Humanities', Miriam Posner's Blog, 27 July 2015, available http://miriamposner.com/blog/whats-next-the-radical-unrealized-potential-of-digital- humanities/>, accessed 4 May 2017.



- 30. Gavan McCarthy and Joanne Evans, 'Mapping the Socio-technical Complexity of Australian Science: From Archival Authorities to Networks of Contextual Information, Journal of Archival Organization, vol. 5, nos. 1-2, January 2008, pp. 149-75, doi:10.1300/J201v05n01 08; Tom J Lynch, 'Social Networks and Archival Context Project: A Case Study of Emerging Cyberinfrastructure', Digital Humanities Quarterly, vol. 8, no. 3, 2014, available at http://www. digitalhumanities.org/dhq/vol/8/3/000184/000184.html>, accessed 28 February 2017; Daniel Pitti, Rachael Hu, Ray Larson, Brian Tingle and Adrian Turner, 'Social Networks and Archival Context: From Project to Cooperative Archival Program', Journal of Archival Organization, vol. 12, nos. 1–2, April 2014, pp. 77–97, doi:10.1080/15332748.2015.999544.
- 31. Joanna Sassoon, 'Sharing Our Story: An Archaeology of Archival Thought. [This Paper Was Originally Presented as a Keynote Address to the Australian Society of Archivists Conference, Alice Springs, August 2007.], Archives and Manuscripts, vol. 35, no. 2, November 2007, p. 45.
- 32. Antonio Cordella, 'Information Infrastructure: An Actor-network Perspective', International Journal of Actor-Network Theory and Technological Innovation, vol. 2, no. 1, 2010, pp. 27-53, doi:10.4018/jantti.2010071602; Fiona Cameron and Helena Robinson, 'Digital Knowledgescapes: Cultural, Theoretical, Practical, and Usage Issues Facing Museum Collection Databases in a Digital Epoch, in Fiona Cameron and Sarah Kenderdine (eds), Theorizing Digital Cultural Heritage: A Critical Discourse, Media in Transition, MIT Press, Cambridge, MA, 2007, pp. 165-91; Sarah Byrne, Anna Clarke, Rodney Harrison and Robin Torrence, 'Networks, Agents and Objects: Frameworks for Unpacking Museum Collections', in Sarah Byrne, Anna Clarke, Rodney Harrison and Robin Torrence (eds), Unpacking the Collection: Networks of Material and Social Agency in the Museum, One World Archaeology, Springer, New York, 2011, pp. 3-26; Ian Hodder, Entangled: An Archaeology of the Relationships between Humans and Things, Wiley-Blackwell, Malden, MA, 2012.
- 33. 'Organic Layout Style', in yFiles for Java Developer's Guide, edition for major version 2.13, 2004, available at http://docs.yworks.com/yfiles/doc/developers-guide/smart_organic_layouter. html>, accessed 28 February 2017.
- 35. See David Berry, 'The Computational Turn: Thinking about the Digital Humanities', Culture Machine, vol. 12, 2011, p. 14.
- 36. Gilmore, p. 109.
- 37. Maud Karpeles, quoted in Stith Thompson (ed.), Four Symposia on Folklore, Indiana University Folklore Series, no. 8, Indiana University Press, Bloomington, IN, pp. 93-4.
- 38. See William Hugh Jansen, 'A Very Preliminary Statement', The Folklore and Folk Music Archivist, vol. 1, no. 1, March 1958, p. 1; Thelma G James, 'Problems of Archives', The Folklore and Folk Music Archivist, vol. 1, no. 2, June 1958, pp. 1, 3.
- 39. Joseph C Hickerson, 'Developing a Folklore Archive', in Richard Mercer Dorson (ed.), Handbook of American Folklore, Indiana University Press, Bloomington, IN, p. 497.
- 40. Dan Ben-Amos, 'Toward a Definition of Folklore in Context', *The Journal of American Folklore*, vol. 84, no. 331, 1971, p. 5, doi:10.2307/539729.
- 41. This emphasis on the importance of the text held in collections and archives is a particular focus of DK Wilgus, 'The Text is the Thing', The Journal of American Folklore, vol. 86, no. 341, 1973, pp. 244-6, doi:10.2307/539153.
- 42. For more on the 'text/context' issue in folklore, see Lisa Gabbert, 'The "Text/Context" Controversy and the Emergence of Behavioral Approaches in Folklore', Folklore Forum, vol. 30, nos. 1/2, 1999, pp. 119-28.
- 43. Timothy R Tangherlini, 'Big Folklore: A Special Issue on Computational Folkloristics', The Journal of American Folklore, vol. 129, no. 511, 2016, pp. 8–9, doi:10.5406/jamerfolk.129.511.0005.
- 44. Source code for the ConneX EAC-CPF visualisation tool is publicly accessible via Bitbucket, available at https://bitbucket.org/esrc/cnex, accessed 4 May 2017.
- 45. Mia Ridge, 'Network Visualisations and the "So What?" Problem', Open Objects, 11 June 2016, available at http://www.openobjects.org.uk/2016/06/network-visualisations-problem/, accessed 18 April 2017.

Acknowledgements

Thanks to Georgia Knight and Christina Giatsios, for their substantial work creating new collection records and enhancing existing data; Nancy Ladas, for her invaluable expertise with EMu and collection data structures; and Richard Gillespie, for his guidance. We also acknowledge Rebe Taylor's significant contribution to the initial McCoy Seed Fund proposal, and thank the McCoy Project for their support.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by the McCoy Seed Fund, an initiative established by the McCoy Project to support the development of collaborative research projects between Museums Victoria and the University of Melbourne.

Notes on contributors

Mike Jones is a PhD candidate in the School of Historical and Philosophical Studies and a consultant research archivist at the eScholarship Research Centre, both at the University of Melbourne. He is also a research associate at Museums Victoria. Since 2008 he has worked as a researcher and consultant on numerous paper-based and digital archival projects, and with academics, government departments, community organisations and the GLAM sector to explore the potential of structured informatic systems to support organisational and public knowledge.

Professor Kate Darian-Smith holds joint appointments as Professor of Australian Studies and History and Chair of the History Program in the School of Historical and Philosophical Studies, and Professor of Cultural Heritage in the Faculty of Architecture, Building and Planning, both at the University of Melbourne. She is an elected Fellow of the Academy of Social Sciences in Australia. Kate is a leader in the fields of history and Australian studies, especially in the areas of war and Australian society, and history and memory.

Deborah Tout-Smith is Senior Curator, Home and Community, at Museums Victoria, working on collections, research and exhibitions. She curates several collections within the Humanities Department of the museum: Home and Community, Clothing and Textiles, Childhood, and the Military History collection. Together these collections comprise over 100,000 objects.

Associate Professor Gavan McCarthy has worked at the University of Melbourne in and around archives since 1978, and is the director of the eScholarship Research Centre. His research interests, predominantly based on action research methodologies, encompass the history and archives of Australian science; contextual information frameworks; archival science and the preservation of knowledge; and the utilisation of network science in social and cultural informatics.

ORCID

Mike Jones http://orcid.org/0000-0002-5448-8799

Kate Darian-Smith http://orcid.org/0000-0001-7773-1205

Gavan McCarthy http://orcid.org/0000-0001-8411-3173