

# digital cultural heritage: FUTURE VISIONS

Edited by Kelly Greenop and Chris Landorf

Papers presented at the *digital cultural heritage: FUTURE VISIONS London Symposium*  
13–15 November 2017 in London, United Kingdom  
<https://www.digitalculturalheritagefutures.com/london-symposium-november-2017>

Burrows, T (2019)

Cultural heritage collections as research data:  
towards computationally amenable collections.

In K Greenop and C Landorf (eds)

*Proceedings of digital cultural heritage: FUTURE VISIONS London Symposium.*

Brisbane:

Centre for Architecture Theory Criticism History,  
pp 1-17.

ISBN-978-0-646-99572-4

© 2017 *digital cultural heritage: FUTURE VISIONS*  
*London Symposium*, London, United Kingdom

Copyright for the proceedings belongs to the Centre for Architecture Theory Criticism History at The University of Queensland. Copyright for the abstracts and papers contained in these proceedings remains the property of the authors. Copyright of images in these proceedings belongs to the authors, or the images appear with permissions granted to those authors. Individual authors were responsible for obtaining the appropriate permission to reproduce any or all images published in the proceedings. The editors and the publisher accept no responsibility if any author has not obtained the appropriate permissions.

Other than fair dealing for the purpose of private study, research or review, as permitted under the Copyright Act, no part of these proceedings may be reproduced by any process without prior permission from the publishers and authors.

The convenors of the *digital cultural heritage: FUTURE VISIONS London Symposium* acknowledge the generous support and sponsorship of the Centre for Architecture Theory Criticism History at The University of Queensland and The Bartlett Real Estate Institute at University College London. Thank you also to the peer reviewers who gave their time and expertise to the refereeing of papers, and Manpreet Dhesi and Carmen Armstrong who provided invaluable project management support.

The papers published in these proceedings are a record of the symposium mentioned on the title page. They reflect the authors' opinions and are published as submitted in their final form with limited editorial change. Their inclusion in this publication does not necessarily constitute endorsement by the editors.

Editors: Kelly Greenop and Chris Landorf

Design and Production: Mark Fletcher

Publisher: Centre for Architecture Theory Criticism History, The University of Queensland

The symposium Convenors received a total of 33 abstracts. All abstracts underwent a double-blind peer review by two members of the Symposium Organising Committee. Authors of accepted abstracts (24) were invited to submit a full paper following presentation of their draft papers at the symposium. All submitted full papers (8) were again double-blind peer reviewed by two anonymous reviewers and given the opportunity to address reviewer comments. Papers were matched as closely as possible to referees in a related field and with similar interests to the authors. Revised papers underwent a final post-symposium review by the editors before notification of acceptance for publication in the symposium proceedings.

Please note that the paper displayed as an abstract only in the proceedings is currently being developed for an edited book on digital cultural heritage.

---

Innovative new data collection and digital visualisation techniques can capture and share historic artefacts, places and practices faster, in greater detail and amongst a wider community than ever before. Creative virtual environments that provide interactive interpretations of place, archives enriched with digital film and audio recordings, histories augmented by crowd-sourced data all have the potential to engage new audiences, engender alternative meanings and enhance current management practices. At a less tangible level, new technologies can also contribute to debates about societal relationships with the historical past, contemporary present and possible futures, as well as drive questions about authenticity, integrity, authorship and the democratisation of heritage.

Yet for many, gaps still exist between these evolving technologies and their application in everyday heritage practice. Following the success of a sister conference in Brisbane, Australia in April 2017, this symposium focused on the emerging disciplines of digital cultural heritage and the established practice of heritage management. The symposium aimed to provide a platform for debate between those developing and applying innovative digital technology, and those seeking to integrate best practice into the preservation, presentation and sustainable management of cultural heritage.

The symposium was designed to encourage critical debate across a wide range of heritage-related disciplines. We welcomed papers from practitioners and academics working in cultural heritage and related fields such as architecture, anthropology, archaeology, geography, media studies, museum studies and tourism. We particularly encouraged papers that explored the challenges of digitising tangible and intangible cultural heritage, those that identified issues with digitisation and digital interaction, and those that addressed the theoretical challenges posed by digital cultural heritage.

**Kelly Greenop and Chris Landorf**  
EDITORS and SYMPOSIUM CONVENORS

---

***digital cultural heritage:***  
***FUTURE VISIONS***  
***London Symposium***

---

Toby Burrows  
University of Oxford  
University of Western Australia

**Abstract:**

*This paper focuses on the re-use of data relating to collections in libraries, museums and archives to address research questions in the humanities. Large-scale research into the history and significance of cultural heritage materials is heavily dependent on the availability of collections data in appropriate formats and on a suitable scale. Until recently, this kind of research has been seriously limited by lack of access to curatorial data. Collection databases have not been available for downloading in their entirety, or have not been made fully available on the Web. There has been a disconnect between curatorial databases and researchers, who have been generally unable to contribute their findings to institutional databases. Some recent “collections as data” initiatives have started to explore approaches to best practice for “computationally amenable collections”, with the aim of “encouraging cultural heritage organizations to develop collections and systems that are more amenable to emerging computational methods and tools’ (Collections as Data 2017)”. This paper discusses three projects that are addressing these issues in similar ways, and uses them to derive lessons and recommendations for future best practice in making collections data available for computational reuse by researchers.*

**Keywords:** Cultural heritage; Collections data; Provenance; Collections data reuse; Computational amenability; Linked dData

# Cultural Heritage Collections as Research Data: Towards Computationally Amenable Collections

This work was supported by the Australian Research Council under Grant LP16010007; the European Union under Grant 626696; and the Economic and Social Research Council (U.K.) under Grant ES/R003971/1.

Dr Toby Burrows is a Senior Researcher in the Oxford e-Research Centre at the University of Oxford, and also holds a Senior Honorary Research Fellowship at the University of Western Australia. He has previously been a Marie Curie International Incoming Fellow at King's College London, and has held visiting fellowships at the University of Pennsylvania, the Vrije Universiteit Amsterdam and Churchill College Cambridge. His research interests include the digital humanities, the history of collecting, and medieval manuscripts.

## Introduction

The importance of cultural heritage collections for research in the humanities, arts and social sciences (HASS) has long been recognised. Though this is not the only reason why such collections have been assembled, it is certainly one of the most important. The recent National Research Infrastructure Roadmap Report for Australia (2017) emphasises the dual nature of these collections; not only do HASS research platforms encompass the physical collections, they also include 'online portals that facilitate the digitisation of and digital access to original artefacts, materials and knowledge' (Australian Government 2017:33). The Report emphasises discoverability and accessibility as priorities, together with 'enhanced digitization, aggregation and interpretation platform processes'. The digital forms of these collections are particularly crucial for research that uses the methodologies, technologies and critical perspectives of the digital humanities (Flanders 2014).

At the same time, an initiative to understand these collections as data is gathering pace in the United States. Under the auspices of the Library of Congress and the Institute of Museum and Library Services, this 'Collections as Data' programme 'aims to foster a strategic approach to developing, describing, providing access to, and encouraging reuse of collections that support computationally-driven research' (Always Already Computational 2017). One of the drivers for this initiative is the perception that, as Miriam Posner argues, 'Libraries and archives [and museums] are increasingly making their materials available online, but, as a general rule, these materials aren't of much use for computational purposes' (Posner 2017).

This paper examines three projects which are addressing the 'collections as data' imperative within the framework of cultural heritage and digital humanities. The first project is 'Collecting the West', in which Western Australian researchers are working with the British Museum to deploy and evaluate the ResearchSpace software, which is designed to integrate heterogeneous collections data into a cultural heritage knowledge graph that can be annotated by researchers. The second project

'Mapping Manuscript Migrations', funded by the Digging into Data programme involves combining collections data from a range of digital sources to reconstruct the histories of large numbers of medieval and Renaissance manuscripts. The third project is HuNI the Humanities Networked Infrastructure, which is building a 'virtual laboratory' for the humanities by reshaping collections data into semantic information networks.

These three projects have been chosen as case studies because they share a common commitment to enabling the reuse of collections data by researchers, and have a similar understanding of the value of collections data as the basis for building knowledge graphs. The lessons learned from the projects will be used to develop recommendations for best practice in the future, presented in the final section of this paper.

### Collections as data

In what sense can collections actually be considered as data? There has been a tendency to try and align digital infrastructure for the humanities and social sciences with the model commonly adopted in the sciences: the data consist of digital content, described by accompanying metadata (Borgman 2007: 215-217). This might work for the social sciences; a service like the Australian Data Archive contains statistical data files, together with descriptive metadata about them. But the analogy begins to break down for the humanities, where, for cultural heritage objects, the descriptions are as important for research purposes as the digital images of cultural heritage objects. The objects themselves, and their digital representations, can only be accessed through the statements that researchers and curators make about them. A more useful approach, it seems to me, is to elide the distinction between 'data' and 'metadata', and to treat descriptions of objects as research data in their own right (Burrows 2011). It follows from this that a productive definition of 'collections data' must encompass the descriptions of objects as well as the objects themselves and their digital representations.

There are various reasons why this kind of collections data is important for research. The most obvious is in order to trace the history of the individual objects

themselves and their relevance as evidence for more general research questions and themes in art, architecture, archaeology, history, literature and other humanities disciplines. As Neil MacGregor vividly demonstrated, the history of almost any object can reveal a great deal about changing cultures over time (MacGregor 2010). A manuscript like the so-called Crusader Bible illustrates this point. Made in mid-thirteenth century France, it later travelled to Naples and then to Krakow, from where it was taken as a gift for the Shah of Persia in Isfahan. Eventually it returned to Western Europe by way of Cairo, and migrated into the world of connoisseurship, manuscript collecting, and conspicuous consumption – ending up in the Morgan Library in New York in the early twentieth century (Abels 2014).

Another important reason for analysing collections data is to investigate the broader history of ownership and collecting. How and why these cultural heritage objects survived to the present day, who has been involved in their history, what they tell us about the priorities and motives of private and institutional collectors alike: these are all important questions underlying the nature of collections today. This speaks to more than just the significance of each specific object; it bears witness to the changing meanings of these objects over time, and to their changing place in a more general social context. And it emphasises the way in which the collections as they exist today are not neutral or objectively representative assemblages of cultural heritage objects. Instead, they reflect the priorities, attitudes and values of specific people and institutions at particular times and places.

A significant project examining these kinds of questions is 'Collecting the West', which is looking at the history of objects relating to Western Australia, many of which are now in British and European collections. An important element in that history involves the acquisition, removal, theft and repatriation of Australian Indigenous artefacts, such as those collected by early European settlers in the 1830s as well as those acquired for re-sale in Europe by travellers like Emile Clement and Paul Denys Montague between the 1890s and the 1930s (Adams 2016).

A project of this type requires the identification and linking of collections data from a range of different sources. In the case of Western Australian Indigenous material, these include the British Museum, the Pitt Rivers Museum in Oxford, the Smithsonian in Washington, and the Hamburg, Frankfurt and Dresden Museums in Europe, as well as numerous collections in Australia and various smaller institutions across the world.

### The Biography of Things

Of particular interest within this type of collections data is the evidence relating to the ownership and provenance of each item. Ownership histories are central to understanding the changing nature of objects over time, and provenance provides crucial evidence for what Igor Kopytoff calls the 'cultural biography of things'. He notes that 'there are many biographies: sheer physical biography, technical biography (repairs), economic biography, social biographies – the owner's economy, ownership and class structure, kin relations' (Kopytoff 1986:68).

Kopytoff is especially interested in the relationship between commoditization, where objects are things which can be bought and sold in the market-place, and singularization, where objects are unique signifiers of cultural value, which exist outside the market-place. As he observes, 'in the homogenised world of commodities, an eventful biography of a thing becomes the story of the various singularisations of it, of classifications and reclassifications in an uncertain world of categories whose importance shifts with every minor change in context' (Kopytoff 1986:89).

For several centuries at least, there has been a thriving market for antiquities, cultural heritage objects and art of many kinds. While the commercial value of specific types of items may have fluctuated significantly over time, it is generally the case that competition to acquire them has been intensifying, to the point where many objects can only be afforded by wealthier individuals and bigger institutions. This has led to such events as the sale of Leonardo's 'Salvator Mundi' for \$450.3 million in 2017 and the sale of the Rothschild Prayer Book for over £8 million in 2014.

This type of 'biography of things in terms of ownership' (Rivers 1910, quoted by Kopytoff) also involves their place in networks of ownership. Ownership does not take place in a vacuum; as Jennifer Van Horn says of 18th-century American private ownership:

Artifacts played an important role in creating cohesion: consumers assembled similar goods to form communities through their shared tastes and distinctive modes of object use. (Van Horn 2017:8).

The collecting and ownership of what we now regard as cultural heritage objects have taken place within networks of shared interests and tastes. As Van Horn points out, there are in fact two networks, or assemblages, involved here,; one of people and one of things:

Networks are often understood as webs that map out a series of interconnected people or, in this case, objects (often artifacts that are related to one another through physical resemblance and common modes of use). (Van Horn 2017: 9).

The ownership of cultural objects at specific times by specific people or organisations tells us something significant about the way in which these objects embody shared cultural values. They show how these networks of ownership change over time, reflecting the changing place of objects in culture and society. In as much as these objects serve as carriers of culture and knowledge, their movements can also reveal the dissemination of ideas across cultures and over time, by a process in which networks of ownership and exchange serve as evidence for networks of knowledge and culture. The best evidence for these patterns can be found in the provenance and ownership history data from collection records.

### Bringing Data data Together

Telling the story of these changing networks of ownership usually means bringing together collections of data from different sources, and from different types of cultural institutions. The technical issues involved in this process are far from trivial. Different

metadata schemas and formats, different vocabularies and different levels of aggregation must all be linked up in a coherent way and exposed through an interface that enables browsing and searching. The most obvious method of doing this is by combining the incoming records into a single database which relies on a standard metadata schema and focuses on the objects themselves. This is the solution preferred by large national and international aggregators such as Trove in Australia, the Digital Public Library of America and even Europeana. But this approach almost inevitably seems to involve reducing the content of each aggregated record to a minimum, affecting both the richness and the discoverability of the data.

More ambitious though more experimental is the use of Linked Data and Semantic Web technologies, focusing on the relationships between objects, persons and events and enabling more complex semantic navigation. This approach has the potential to retain much more of the semantic richness of the data, and even to add value to it by situating it within a broader context of knowledge graphs and networks (Hyvönen 2012). Some projects using this approach are discussed below.

Political and policy issues are equally significant. Institutions vary greatly in their willingness to share data, for several different reasons: a feeling that their data are not of sufficient quality; an assumption that data are important intellectual property; a need to raise revenue through sales of digitized objects; and so on. They also vary greatly in their ability to share data (from a technical point of view), or to support all but the simplest type of export or download. In relation to Linked Data specifically, the institution may well take the view (as the National Library of Australia has done recently) that the 'business case' has not yet been proved – that is, that there is insufficient demand to justify investing time and money in establishing suitable processes.

### Collections Data in Action

Despite this range of issues and potential barriers, there are a growing number of examples of major institutions sharing their collections data, in both the

museum and library worlds. In the library sector, the OPenn service makes available descriptive data and images relating to manuscripts held in the University of Pennsylvania Library. Each manuscript has a descriptive file (TEI-encoded<sup>1</sup>) and a set of digital images, all of which can be downloaded and reused freely. The Bodleian Library (Oxford University) is doing something similar for its new medieval manuscripts catalogue. In addition to a new Web catalogue, the descriptive data are available for download from a GitHub<sup>2</sup> repository as TEI-encoded files.

Several major museums have also made their collections data available, enabling researchers and users to analyse, create and play (Fitzpatrick 2017). The release of the collections data of the Museum of Modern Art (MoMA) in 2014 led to a series of experiments, including an analysis of the collection of paintings by size and an analysis of the Museum's acquisition activities in which year of creation was mapped against year of acquisition. The Tate Gallery (London) released its collections data in 2013 as CSV files.<sup>3</sup> Among the uses made of the data was to re-work them as a network graph using the Neo4j<sup>4</sup> software (Cunningham 2014). This enabled users to find and display (among other things) the shortest path of relationships between two artists, such as Augustus John and William Johnstone.

More unexpected and entertaining uses of the collections data have included Twitter bots which automatically tweet database records from institutions like the Tate Gallery and the Rijksmuseum (Amsterdam), accompanied by an image of the object. More unusual still was a performance of MoMA's (New York) collections data as a series of spoken texts read by Museum staff (Thorpe 2015).

Several current projects are working with collections data to answer complex research questions and build humanities-oriented infrastructure. The 'Collecting the West' project is bringing together data relating to Western Australian objects held in collections in Australia and Europe. The software being used is the British Museum's ResearchSpace, which maps collections data to the CIDOC-CRM<sup>5</sup> ontology and enables complex semantic exploration of the



---

results. Initially limited to the British Museum's own collections, it is now been tested by other institutions in Europe and North America. ResearchSpace, which is built from the Metaphactory platform, enables researchers to work with collections data by adding annotations and arguments to objects and other entities.

OxLOD (Oxford Linked Open Data) is taking a generally similar approach, mapping heterogeneous collections data from Oxford University's museums and libraries to produce an interdisciplinary platform for cultural heritage research. An estimated 200,000 digital records will be linked and mapped in the initial phase of this project. This project builds on work previously done for the CLAROS initiative, and is employing a Linked Data approach based on the CIDOC-CRM ontology. Oxford University did conducted a survey of its cultural heritage collections in 2016, which identified significant gaps in descriptive data about the collections, with about 40% undescribed (Cannon and Madsen 2016). It also revealed that only a tiny proportion of the collections have been digitised – probably less than 1%.

Another project working in the same general area is 'Mapping Manuscript Migrations', funded by the Digging into Data Challenge for 2017-2019. A collaboration between institutions in the United Kingdom, the United States France and Finland, this project is bringing together data from a range of major databases to map and analyse the histories of as many manuscripts as possible. The initial data sources include the Schoenberg Database of Manuscripts, which contains observations relating to the provenance of specific manuscripts; the Bodleian Library's new Medieval Manuscripts Catalogue, which contains TEI-encoded descriptions; and two databases services from the Institut de recherche et d'histoire des textes – the Medium database, with brief descriptions of manuscripts, and the Bibale database, with provenance records. Data from these sources are linked through a common data model, and expressed as Linked Data in RDF<sup>6</sup> format.

This project in its turn builds on an earlier investigation into the history of the vast manuscript collection of

---

Sir Thomas Phillipps (1792-1872). Using collections data from various library and museum sources, I traced the history of a sample of his more than 40,000 manuscripts, which were dispersed to a range of public and institutional collections in the century after his death. The histories of individual manuscripts can be mapped and visualised, together with the network graph of people, places and institutions involved in these events (Burrows 2017). Phillipps collected all kinds of manuscripts, from beautiful and lavish volumes to ephemeral scraps of paper, and the history of his collection reveals a good deal about the interplay between connoisseurship and antiquarianism. The evidence for his activities is large and varied, but there are various difficulties with making use of the data. Many modern collections are poorly documented, with objects lacking any kind of descriptive data. Many library and museum databases handle provenance information in a way that is difficult to use computationally, and a surprising number make it very hard to download bulk data.

Another service which relies heavily on the reuse of collections data is HuNI, the Australian virtual laboratory for the humanities, which ingests records from library catalogues as well as data from various archives (Burrows and Verhoeven 2015). It also aggregates data from the Trove digitised newspaper collection, and from reference works, bibliographies and event-oriented databases, amounting to more than thirty in all. HuNI has recently added a pipeline from ingesting data for collections created with the Omeka<sup>7</sup> software.

HuNI re-formats collections data by extracting entities from incoming records and making them available for linking and visualising, in the form of nodes on a network graph. Interpretations can be added to the data by users, in the form of relationships and links between nodes, using terminology created by the user. Entities can also be selected and saved in users' own collections, employing their own categorisations and classifications. The network can be explored visually, and can be searched for all nodes connected to a specific node, up to five links away, as well as for the shortest path between two nodes.

---

One of HuNI's main aims is to enable vernacular, user-driven knowledge structures, rather than simply importing those found in the data sources. This is one of the biggest issues arising from the reuse of collections data: the way in which these databases embody a specific set of terminologies, vocabularies and ontologies – a particular view of the world. Whose authority do the data represent? For the most part, this is the curatorial perspective of the collecting institutions, rather than the perspectives of researchers or of the wider community – let alone the perspectives of the indigenous communities from which many significant cultural objects originate.

What these projects and services have in common is the idea of taking collections data and using them to create new knowledge structures, especially in the form of network graphs of relationships between entities – including people, places and objects. While there are other things that can be done with collections data (such as the comparison of images using the International Image Interoperability Framework,<sup>8</sup> and the analysis of the texts carried by cultural objects), network graphs are a powerful way of uncovering the meaning and significance of the knowledge embedded in cultural heritage collections.

These three projects share a common interest in using collections data to answer research questions, not just as a route to discovering the contents of cultural heritage collections. They aim to make it possible for research to work actively with collections data, rather than simply consuming collections data for searching and browsing. They also show how collections data can be exploited to address research questions around the nature of collections themselves and around the development of knowledge graphs.

### Next Steps

To make services like these possible, collections data need to be made available in certain ways and under certain conditions. Recommendations for best practice, at the moment, tend to be focused mostly on processes and procedures, encompassing download formats, licensing, and availability in particular (Fitzpatrick 2017). These are undoubtedly important;

---

having collections data easily accessible in bulk on the Web, under a Creative Commons licence that permits free reuse, is essential. Download formats are more debatable: APIs<sup>9</sup> are not necessarily the best approach, given that their use is likely to require a significant level of technical expertise (Tauberer 2014). XML<sup>10</sup> dumps and CSV files are easier to use, but may not contain all the elements in the source database.

As the interest of researchers in reusing collections data continues to grow, however, cultural heritage institutions increasingly need to start looking beyond simply making their data available for bulk downloading or via an API. One of the major use cases is to link together data from different institutions, without diminishing the semantic richness, in order to ask questions on a larger scale. At the moment, researchers are having to do much of this work themselves. This raises two important questions: should institutions help this process, and what kind of infrastructure might be built as a result?

The prominence of Linked Data in the solutions being adopted by researchers strongly suggests that institutions should make their data available in formats suitable for incorporation into Linked Data environments. While many institutions might not yet see a 'business case' for this approach, others like the British Library and the British Museum have already followed this route. Making available an RDF version of a relational database would be a significant contribution. But even embedding into that database identifiers that point to widely-used Linked Data ontologies and vocabularies like VIAF,<sup>11</sup> GeoNames<sup>12</sup> and Wikidata<sup>13</sup> would be valuable. So too would taking a critical look at ways of improving the computational value of ownership and provenance data in these records. Enabling researchers and curators to annotate and add to the data is also emerging as an important requirement.

Beyond this, though, lies the wider landscape of digital infrastructure. The Santa Barbara Statement on Collections as Data (2017) observes that 'Working toward interoperability entails alignment with emerging and/or established community standards and infrastructure.' At present, the Linked Data

landscape is largely being built by research groups rather than cultural institutions, which still tend to focus on their own collections. In this context, an initiative like 'Linked Pasts', which has emerged from the Pelagios Commons, is an important development, offering a vision of joining up disparate Linked Data projects in the humanities to create a 'wider ecosystem' (Grossner and Hill 2017).

As long as these kinds of initiatives remain tied to research projects, their future sustainability will be reliant on the uncertainty of grant funding. Collecting institutions should look closely at them as outcomes of the reuse of collections data, and consider seriously the value of partnerships with the researchers involved. Building knowledge networks that represent the history and transmission of culture as seen through the biographies of objects is a major research goal. Collections data have a vital role to play in that process.

#### Acknowledgements

This work was supported by the Australian Research Council under Grant LP16010007; the European Union under Grant 626696; and the Economic and Social Research Council (U.K.) under Grant ES/R003971/1.

#### References

- Abels, R. 2014. "Cultural Representations of Warfare in the High Middle Ages: The Morgan Picture Bible." In *Crusading and Warfare in the Middle Ages: Realities and Representations: Essays in Honour of John France*, edited by S. John and N. Morton, 13-35. Farnham: Ashgate.
- Adams, J. 2016. "Magic and Memory: Paul Denys Montague's Collection from New Caledonia." *Journal of the History of Collections* 28 (2): 279-292.
- Always Already Computational: Library Collections as Data*. 2017. <https://collectionsasdata.github.io/>
- Australian Government. 2016. *National Research Infrastructure Roadmap*. Canberra: Australian Government.
- Borgman, C. L. 2007. *Scholarship in the Digital Age: Information, Infrastructure, and the Internet*. Cambridge: MIT Press.
- Burrows, T. 2011. "Sharing Humanities Data for E-Research: Conceptual and Technical Issues." In *Humanities Perspectives on Digital Scholarship*, edited by N. Thieberger, 177-192. Melbourne: PARADISEC. <http://ses.library.usyd.edu.au/bitstream/2123/7938/1/Burrows.pdf>
- Burrows, T. 2017. "The History and Provenance of Manuscripts in the Collection of Sir Thomas Phillipps: New Approaches to Digital Representation." *Speculum* 92 S1:S39-S64.
- Burrows, T., and D. Verhoeven. 2015. "Aggregating Cultural Heritage Data for Research Use: The Humanities Networked Infrastructure (HuNI)." In *Metadata and Semantics Research, 9th Research Conference, MTSR 2015, Manchester, UK, September 9-11, 2015: Proceedings*, edited by E. Garoufallou, R. J. Hartley, and P. Gaitanou, 417-423. Cham: Springer.
- Cannon, C., and C. Madsen. 2016. "The Future of Finding at the University of Oxford." Paper presented at CNI Fall Membership Meeting, Washington D.C. Accessed 12-13 December 2016. <https://www.slideshare.net/mccarthymadsen>
- Collections as Data. 2017. "Shaping Humanities Data: Use, Reuse, and Paths Toward Computationally Amenable Cultural Heritage Collections." <https://collectionsasdata.github.io/dh2017/>
- Cunningham, L. 2014. "Graphing the Tate." *Larkin Cunningham* [blog]. Accessed 22 December 2014. <http://larkin.io/index.php/2014/12/22/graphing-the-tate/>
- Fitzpatrick, L. K. 2017. "Shared practices in museum open collections data." *Medium*. Accessed 22 February 2017. <https://medium.com/berkman-klein-center/shared-practices-in-museum-open-collections-data-72e924c4849a>
- Flanders, J. 2014. "Rethinking Collections." In *Advancing the Digital Humanities*, edited by P. Longley Arthur and K. Bode, 163-174. London: Palgrave MacMillan.
- Grossner, K., and T. Hill. 2017. "From linking places to a linked pasts network." [http://kgeographer.com/pubs/LinkedPastsNetwork\\_7Dec.pdf](http://kgeographer.com/pubs/LinkedPastsNetwork_7Dec.pdf)
- Hyvönen, Eero. 2012. *Publishing and Using Cultural Heritage Linked Data on the Semantic Web*. San Rafael, CA: Morgan & Claypool.

Kopytoff, I. 1986. "The Cultural Biography of Things." In *The social life of things*, edited by A. Appadurai, 64-91. Cambridge: Cambridge University Press.

MacGregor, N. 2010. *A History of the World in 100 Objects*. London: Penguin.

Posner, M. 2017. "Actually useful collection data: some infrastructure suggestions." In *Always Already Computational: Library Collections as Data: National Forum Position Statements*. [https://github.com/collectionsasdata/collectionsasdata.github.io/raw/master/aac\\_positionstatements.pdf](https://github.com/collectionsasdata/collectionsasdata.github.io/raw/master/aac_positionstatements.pdf)

Rivers, W. H. L. 1910. "The Genealogical Method of Anthropological Inquiry." *Sociological Review* 3 (3): 1-12. *Santa Barbara Statement on Collections as Data*. 2017. <https://collectionsasdata.github.io/statement/>

Tauberer, J. 2014. "Bulk Data or an API?" <https://opengovdata.io/2014/bulk-data-an-api/>

Thorp, J. 2015. "A sort of joy: performing MoMA's 120,000 object collections database." *Medium*. Accessed 27 July 2015. <https://medium.com/memo-random/a-sort-of-joy-1d9d5ff02ac9>

Van Horn, J. 2017. *The Power of Objects in Eighteenth-Century America*. Chapel Hill: University of North Carolina Press.

## Endnotes

- 1 The Text Encoding Initiative (TEI) is "a consortium which collectively develops and maintains a standard for the representation of texts in digital form. Its chief deliverable is a set of Guidelines which specify encoding methods for machine-readable texts, chiefly in the humanities, social sciences and linguistics." See <http://www.tei-c.org/>
- 2 GitHub "is a web-based version-control and collaboration platform for software developers... Git is used to store the source code for a project and track the complete history of all changes to that code." (<https://searchitoperations.techtarget.com/definition/GitHub>)
- 3 CSV files are comma separated values files, that can be imported into any spreadsheet or relational database software.
- 4 Neo4j is an open-source software for managing graph databases.
- 5 CIDOC-CRM is the Conceptual Reference Model (CRM) of the Comité International pour la Documentation (CIDOC), in English the International Committee for Documentation, of cultural heritage implemented by ICOM, the International Council of Museums. It "provides definitions and a formal structure for describing the implicit and explicit concepts and relationships used in cultural heritage documentation" see <http://www.cidoc-crm.org/>
- 6 Resource Description Framework (RDF) "is a standard model for data interchange on the Web. RDF has features that facilitate data merging even if the underlying schemas differ, and it specifically supports the evolution of schemas over time without requiring all the data consumers to be changed" see <https://www.w3.org/RDF/>
- 7 Omeka software "provides open-source web publishing platforms for sharing digital collections and creating media-rich online exhibits" see <https://omeka.org/>
- 8 International Image Interoperability Framework (IIIF) provides common application programming interfaces that support interoperability between image repositories, to enable ease of viewing both images and their associated metadata, see <http://iiif.io/about/>
- 9 API is an Application Programming Interface, a "set of commands, functions, protocols, and objects that programmers can use to create software or interact with an external system" see <https://techterms.com/definition/api>
- 10 Extensible Markup Language (XML) is "is a simple text-based format for representing structured information" that is both human-readable and machine-readable, see <https://www.w3.org/standards/xml/core>
- 11 VIAF (Virtual International Authority File) "combines multiple name authority files into a single OCLC-hosted name authority service. The goal of the service is to lower the cost and increase the utility of library authority files by matching and linking widely-used authority files and making that information available on the Web" see <https://viaf.org/>
- 12 GeoNames is a "geographical database covers all countries and contains over eleven million placenames that are available for download free of charge" see <http://www.geonames.org/>
- 13 Wikidata "is a free and open knowledge base that can be read and edited by both humans and machines. Wikidata acts as central storage for the structured data of its Wikimedia sister projects including Wikipedia, Wikivoyage, Wikisource, and others" see <https://www.wikidata.org>



THE UNIVERSITY  
OF QUEENSLAND  
AUSTRALIA

**school of  
architecture**

**a | t | c | h**

Architecture  
Theory  
Criticism  
History