

Representing a Relationship:

The use of metadata to represent the relationship between physical
objects and their digital surrogates

By

Laura Deborah Jamieson

A thesis submitted to the Victoria University of Wellington
in fulfilment of the requirements for the degree of
Master of Arts in Museum and Heritage Studies

Victoria University of Wellington

2020

Abstract

Digital collections are increasingly prominent in museums as born-digital material is acquired by institutions, and digital surrogates of physical items are created through digital imaging, digitisation, and reformatting projects. These digital collections are a significant development in museums and a useful tool, particularly for access. When a digital surrogate is created of a physical object, they have an inherent connection to one another. Representing this relationship is important for museums in order to provide context for their collection items. These types of relationships also occur across physical formats, and the consequence of a breakdown in this relationship has been shown in the literature to lead to a loss of context. However, it is unclear how the relationship a physical object has with its digital surrogate is represented in the metadata. Current literature on digital collections only briefly explores existing relationships between digital and physical collections and provides no framework for best practice in a museum context.

This thesis examines how metadata is used to represent the relationship between a physical object and its digital surrogate at the Museum of New Zealand Te Papa Tongarewa. The research involved a single-site case study, with interviews and documentary research which were thematically analysed. This thesis shows how the relationship between physical and digital objects are primarily represented at Te Papa through the collection management system's structure, with some metadata elements representing the relationship incidentally. It also shows that there are differing worldviews and perspectives across the GLAM domains in the language and the drivers of digitisation.

This research serves as a snapshot of current practice at one institution and encourages further research to better understand the long-term implications of this and other approaches. For museums, understanding how the relationship between physical objects and digital surrogates is currently being represented through metadata could help support professional practice for both types of collections, ensure the relationship is maintained, and help support existing and future digital interventions in museums.

Acknowledgements

I wish to acknowledge all those who have provided support and guidance during this process.

In particular, I would like to thank my supervisors, Dr Lee Davidson and Dr Shannon Wellington, who have encouraged and guided me through the thesis with great care.

I wish to acknowledge the support I received as the recipient of the inaugural Te Papa Dame Cheryll Sotheran Memorial Scholarship, and the Victoria Master's by Thesis Scholarship. Special thanks to the family of Cheryll Sotheran, Te Papa, and Victoria University of Wellington; I could not have undertaken this thesis without this support.

My thanks also go to those at Te Papa who were involved in the research, particularly my interview participants.

Special thanks to Logan, my family, fellow postgraduates, and friends who have all provided support, distraction, and love throughout this process.

A professional editor, Bella Mae, provided proof-reading services in accordance with the Editorial Advice Policy of Victoria University of Wellington

Table of Contents

Abstract	i
Acknowledgements	ii
Table of Contents	iii
List of figures and tables	v
Introduction	1
Literature review	4
Research questions	20
Research design	23
Thesis structure	34
Chapter one: Case study - Museum of New Zealand, Te Papa Tongarewa	35
Introduction to Te Papa	35
Digitisation and collection management at Te Papa	37
Interview participants	43
Chapter two: Different world views, perspectives, definitions, and domains	45
Different world views and perspectives	45
Digitisation for access and preservation	51
Different domains	54
Summary	56
Chapter three: System architecture representing relationships	59
System structure and relationships	59
Technical and descriptive metadata	64
Collections Online and fuzzy search	66
Summary	68

Chapter four: Types of metadata representing the relationship	69
Metadata and object relationships	69
Examples of metadata representing the relationship	75
Summary	79
Conclusion	81
Bibliography	89
Appendices.....	102
Appendix one: Interview guide.....	102
Appendix two: Codes and themes.....	105

List of figures and tables

Table 1:	Documentary sources	31
Figure 1:	Image of catalogue record in EMu (IRN 1347076)	39
Figure 2:	Image of media asset record in EMu (IRN 1601658)	40
Figure 3:	Image of multimedia record in EMu (IRN 400818)	41
Figure 4:	Image of page two of search results on Collections Online, Te Papa for 'Rita Angus'	42
Figure 5:	EMu record structure	60
Figure 6:	Collection object catalogue record (IRN 22493)	61
Figure 7:	EMu record structure showing the hierarchy of records	64
Figure 8:	Image of 'Collection Objects' on Rita Angus Person page	71
Figure 8a:	(insert) Expanded menu showing filter options	71

Introduction

When a digital surrogate is created of a collection object, the physical object and the digital asset that is created have a fundamental relationship with one another.

Representing this using metadata is one key way to support the connection between these two over time and provide context to both items. Digital surrogates, the outputs of digitisation work, are being created by collecting institutions, including libraries, archives, and museums, at a rapid rate. In New Zealand, two of the large museums have recently completed, or are currently in the midst of rapid digitisation projects, outside of their business as usual digitisation work.

The relationship between the object and digital surrogate is important to retain in order to help provide context and information about both the object and the new digital asset. Equally, retaining this relationship can help support the drivers of digitisation, access and preservation. Metadata, the information, that surrounds both the collection object and the digital surrogate represents the relationships they have with other objects, ideas, and entities, and for this reason, would be a likely and obvious place for information to be found which represents the specific relationship between a digital surrogate and the object it was created from. For this reason, this thesis explores how metadata is used to represent the relationship between physical collection objects and their digital surrogates at the Museum of New Zealand Te Papa Tongarewa (Te Papa).

This chapter looks at the context for this research before discussing the literature on digitisation, digital collections, and collection management, and seeking out areas of connection and disconnection in the relationships between digital and physical collections. Definitions for the specialised language used in the research are provided before the research questions are then introduced. The research methodology is then explored. The introduction concludes with the thesis structure.

Research context

Digital collections encompass born-digital material, as well as reformatted analogue materials and digital surrogates of physical collection items. My previous work with documentary heritage collections has given me experience in the creation and care of digital collections, and in the intersection of digital collections with the physical collection items. Conversations around digital surrogates, and whether these required separate records or numbering, were of particular interest. I was also involved with the different teams dealing with each collection, each using different processes which could cause issues. I was also able to see first-hand the ways that users of the material interacted with each format.

I was interested in conducting research which supports and strengthens professional practice, as well as looking at digital collections from a point of view that is grounded in collection management. Looking at born-digital collections was considered, but these do not centre on the relationship with physical collections in the same way as digital surrogates. While this is an important area to explore, focusing on digital surrogates allowed me to look at both digital and analogue collection formats, whilst keeping the idea of the relationship between the two central.

The original idea for this research covered this broad area and spanned both documentary heritage and museum collections. This was narrowed down to the research question, which focuses in on a specific area of inquiry and looks at a single site case study in order to get a small snapshot of current practice at a large institution. A key part of this process was being specific about what I was interested in and would be able to achieve, as well as exploring what would be most useful to find out about.

Focusing on the role of metadata, the information which surrounds collection objects, provided the opportunity to think about how it functions. By looking at how it represents the relationship, rather than how it 'maintains' it, implies the relationship

and describes the role of the metadata, without adding an ongoing agency that may not exist.

Further exploration into dealing with the intersection of physical collections with digital collections interested me, due to the variety of ways this has been approached, but also because of the lack of literature and best practice in this area in museums. The literature was reviewed to look for areas of connection and disconnection between physical and digital collections, with a focus on the relationship and collection management.

Literature review

Introduction

This literature review covers the existing writing on digital collections, which includes digital surrogates, and examines whether this intersects with the literature on physical collections in order to discover how the relationship between the two collection formats is discussed and any areas of connection or disconnection. Starting with the history and context of the field, it looks at digital collections, digital surrogates, and digitisation, as well as metadata, and explores ideas of context and relationships before picking up on other concerns in the literature on digital collections including authenticity and resource. This review aims to draw out whether the literature talks about digital and physical collections separately or makes connections between them. I talk about connection and disconnection throughout to highlight the ways in which the relationship between the two types of collections is talked about in the literature. Ultimately the relationship between digital and physical collections is touched on across the literature but there is a lack of research which focuses on the relationship between the two types of collections, as well as what the consequences of both connection and disconnection for museum collections might be.

Within the cultural heritage field, literature is generally split along the domains of museums, archives, and libraries. While there is some overlap between them, concepts and ideas tend to be explored within these domains, rather than across them. However, collection management, digital surrogates, and metadata are areas of exploration in all of these domains. As I am specifically looking at the collecting model that is common to museums, this forms the main portion of literature consulted. Alongside this, I have also consulted some archives literature particularly relating to metadata and collection relationships. Some libraries literature has been consulted, mainly where it deals with libraries which have documentary heritage reference collections. This excludes literature which deals with public library collections which is outside the scope of my research.

History and context

While plenty has been written about caring for physical collections, covering management, care, preservation and conservation, and best practice in all areas, these elements as they relate to digital collections are less prevalent in museums literature. Digital collections have been part of memory institutions for four decades but only more recently have become a common and expected part of museum practice.¹ The uptake of digitisation projects in the 1980s occurred throughout the GLAM (Galleries, Libraries, Archives, and Museums) sector, but the focus on documentary heritage and the common formats across libraries and archives has seen a more universal uptake of digitisation projects and an earlier intake of born-digital material.² Libraries and archives literature reflects this early commitment to experimentation and universal uptake with a higher volume of literature exploring digital collections than seen in the literature on museums and galleries.³ Looking across the domains adds value to the museums literature that is my focus.

Digital collection, digital surrogates, and relationships

Digital collections have a variety of definitions attached to them, to expand or reduce what is covered by the term. Commonly included are born-digital material and digital surrogates of physical collection objects. Often digitally stored records and metadata of digital and physical collection objects are part of the definition as well. Digital collections are also written about as part of a broader discussion of digital inclusion in museums, described by Henning as ‘New Media’.⁴ Here digital collections, and the activities related to them - including online collection access and databases, are included in this broader digital category. Multi-format or hybrid collections are also now common, as groups of items which are in multiple formats are acquired by institutions. Helfrich notes that as the world transitions from analogue to digital, so

¹ Melissa M. Terras, ‘The Rise of Digitization’, in *Digitisation Perspectives*, ed. Ruth Rikowski, Educational Futures Rethinking Theory and Practice (Rotterdam: Sense Publishers, 2011), 3–4.

² Terras, ‘The Rise of Digitization’, 3–10.

³ Terras, ‘The Rise of Digitization’, 16.

⁴ Michelle Henning, ‘New Media’, in *Companion to Museum Studies*, ed. Sharon Macdonald (Wiley, 2008).

do the items that are acquired, and therefore we could consider that all archival collections in the future will be hybrid collections.⁵

Digital surrogates, in particular, have an inherent relationship with the physical object from which they were created, regardless of whether that remains clear in the record or metadata. Conway has suggested that the relationship that a digital surrogate has with its source material “conforms to the ‘law of contact’”, which suggests that things which have had contact continue to act upon each other after this physical contact has been broken.⁶ An absence of anything that represents this could lead to disconnection. This is also reflected in the literature consulted, which does not refer back to physical collections when discussing digital surrogates. Once physical collections have been reformatted, digitised, or have undergone digital imaging, the literature shifts focus to ensuring the preservation of this new digital collection. Records and metadata emerge as a key area to ascertain whether disconnection of the relationship has occurred, particularly in the libraries literature.⁷ The *IFLA/UNESCO Survey on Digitisation and Preservation* found a mixed response when asking whether institutions created a separate record for a digitised item.⁸ In *Born Digital in New Zealand: Report of Survey Results* they estimate from the data they collected that “at least half of born-digital material held by institutions is not described and not under sufficient intellectual control”.⁹ This suggests that the metadata and records for both digital surrogates and born-digital material may be insufficient for maintaining relationships with physical collections, and so leads to disconnection.

⁵ Kurt G. F. Helfrich, ‘Questions of Authenticity: Challenges in Archiving Born-Digital Design Records’, *Art Libraries Journal* 35, no. 3 (2010): 23.

⁶ Paul Conway, ‘Digital Transformations and the Archival Nature of Surrogates’, *Archival Science* 15, no. 1 (2014): 53.

⁷ Joyce Ray, ‘The Rise of Digital Curation and Cyberinfrastructure’, *Library Hi Tech* 30, no. 4 (2012): 611.

⁸ Sara Gould et al., *IFLA/UNESCO Survey on Digitisation and Preservation*, International Preservation Issues; No. 2 (Wetherby, UK; Paris, France: IFLA-UAP; IFLA-PAC, 1999); 22.

⁹ Jessica Moran, ‘Born Digital in New Zealand: Report of Survey Results’ (Wellington, New Zealand: Alexander Turnbull Library, National Library of New Zealand, 2017), 13.

Context, relationality, and the role of metadata

The relationships between objects are important because they provide context to the items. Context is a common idea in libraries literature. Beaudoin defines this context as “those properties of an object related to its creation and preservation that make the object’s origin, composition, and purpose clear” and notes that recording the context is of particular importance for physical objects that are digitally preserved and non-textual objects.¹⁰ Even simple characteristics of a physical object, such as size and scale, can be lost in the digital surrogate if not recorded.¹¹ Of course, context has broad meanings across environments, however, most are encompassed by “the interrelated conditions in which something exists or occurs”.¹² In this way, context can be said to directly relate to relationships. Unsworth notes that digital materials pose “... the risk of decontextualization — the possibility that the digital surrogate will become detached from some element of context that is important to understanding what it is, and will be received and understood in the absence of that context”.¹³ Equally, Lee says that “relationships to other digital objects can dramatically affect the ways in which digital objects have been perceived and experienced”.¹⁴

When discussing the relationships between types of collections, and collection items, the museums literature often refers to the concept of relationality. This is the interconnectedness of things, “the quality of always already being related; it is through relational entanglement that the characteristics and meanings of things emerge”.¹⁵ Fiona Cameron puts forward that our current documentation structures

¹⁰ Joan E. Beaudoin, ‘Context and Its Role in the Digital Preservation of Cultural Objects’, *D-Lib Magazine* 18, no. 11/12 (November 2012): 2; Joan E. Beaudoin, ‘A Framework for Contextual Metadata Used in the Digital Preservation of Cultural Objects’, *D-Lib Magazine* 18, no. 11/12 (November 2012): 3.

¹¹ Beaudoin, ‘Context and Its Role in the Digital Preservation of Cultural Objects’, 10.

¹² Holger Brocks et al., ‘Modeling Context for Digital Preservation’, in *Smart Information and Knowledge Management: Advances, Challenges, and Critical Issues*, ed. Edward Szczerbicki and Ngoc Thanh Nguyen, Studies in Computational Intelligence (Berlin, Heidelberg: Springer Berlin Heidelberg, 2010), 197.

¹³ John Unsworth, ‘The Value of Digitization for Libraries and Humanities Scholarship’ (An Innodata Isogen Symposium, The Newberry Library, 2004), accessed 11 September 2018, <http://www.people.virginia.edu/~jmu2m/newberry.04.html>.

¹⁴ Christopher A. Lee, ‘A Framework for Contextual Information in Digital Collections’, *Journal of Documentation* 67, no. 1 (2011): 100.

¹⁵ Michael Alastair Jones, ‘Documenting Artefacts and Archives in the Relational Museum’ (PhD, The University of Melbourne, 2018), 8.

should be adjusted to better contextualise collections.¹⁶ Of particular relevance is the focus on digital documentation, including imaging, at acquisition to better link existing documentation held across different spaces.¹⁷ Cameron is interested in ensuring that collection items are not unnecessarily tied to the classifications that the institutional context puts them into.¹⁸ This will ensure knowledge frameworks sit more broadly, across collections and institutions. Jones points out that collections often use words which allude to this sense of relationality, and it is difficult to talk about a discrete object, without exploring how it is interconnected.¹⁹ As Besser points out, “in the digital world, information is increasingly inter-related to other information”.²⁰ The relationships between objects and information are highlighted as of importance to museum collections, especially with digital systems.

Archival arrangement and description prioritises the relationships between items and context, in order to preserve evidentiary value.²¹ The archival bond brings together all items created as part of the same activity; each item is an object and so the archival bond allows them to retain their meaning.²² Duranti expands on this with the idea of an archival bond, which is distinguished from a more general context; context is outside of the record, whereas the archival bond is an essential part of the record.²³ This approach seems to support the approach to relationality discussed in the museums literature. For item level systems, such as those often used for museum collections, this suggests a greater need to capture context, but also relationships, as this becomes the key area to provide meaning to and interpretation of the object. For

¹⁶ Fiona Cameron, ‘Museum Collections, Documentation, and Shifting Knowledge Paradigms’, in *Reinventing the Museum: The Evolving Conversation on the Paradigm Shift*, ed. Gail Anderson, 2nd ed. (Lanham, Md.: AltaMira Press, 2012), 224–227.

¹⁷ Cameron, ‘Museum Collections, Documentation, and Shifting Knowledge Paradigms’, 233.

¹⁸ Cameron, ‘Museum Collections, Documentation, and Shifting Knowledge Paradigms’, 227.

¹⁹ Jones, ‘Documenting Artefacts and Archives in the Relational Museum’, 273.

²⁰ Howard Besser, ‘Digital Longevity’, in *Handbook For Digital Projects: A Management Tool for Preservation and Access*, ed. Maxine K. Sitts, First Edition (Massachusetts, USA: Northeast Document Conservation Center, 2000), 168.

²¹ Jeff Crow et al., ‘A Unique Arrangement: Organizing Collections for Digital Libraries, Archives, and Repositories’, in *Theory and Practice of Digital Libraries*, ed. Panayiotis Zaphiris et al., Lecture Notes in Computer Science (Springer Berlin Heidelberg, 2012), 335–356.

²² Crow et al., ‘A Unique Arrangement’, 336.

²³ Luciana Duranti, ‘The Archival Bond’, *Archives and Museum Informatics* 11, no. 3 (1 September 1997): 217.

digital surrogates, their relationship with the object they were created from provides their evidentiary value.

There are, however, clear issues in implementing and maintaining the relationality between different types of collections within museums. In particular, Jones explores how collection items could benefit from information found in field notebooks which are categorised as part of the museum archive.²⁴ These connected items have become disconnected within the museum environment and a move towards relationality could improve this relationship. Blackaby and Sandore note that “free movement ought to be enabled from the contextual materials to the collection items”.²⁵

These concerns about loss of context often find their solutions in libraries literature in the recording of metadata to sit alongside the physical and digital objects. Metadata functions as a way of creating order and capturing information, and when associated with resources can be embedded, such as in mark-up; associated, in records; or third-party, external, or separate in some way.²⁶ Beaudoin notes eight key areas of context that should be captured in metadata to provide this context, showing the range of context that can be captured and used to represent the relationship that objects have to one another, particularly surrogates.²⁷ Brocks et al. also suggests that attributes of digital objects and the relations between them and collections are important to represent, as well as “the processes in which they were created, preserved, accessed and reused”.²⁸ Lee has also suggested a framework which takes a target digital object and relates it to contextual entities in order to formalise the types of relationships and contexts a digital object may have, including over time.²⁹ The number of frameworks and schema put forward in the literature suggests a problem

²⁴ Jones, ‘Documenting Artefacts and Archives in the Relational Museum’, Chapter 3.

²⁵ Jim Blackaby and Beth Sandore, ‘Building Integrated Museum Information Retrieval Systems: Practical Approaches to Data Organization and Access’, *Archives and Museum Informatics* 11 (1997): 144.

²⁶ Erik Duval et al., ‘Metadata Principles and Practicalities’, *D-Lib Magazine* 8, no. 4 (April 2002), 10, 15.

²⁷ Beaudoin, ‘A Framework for Contextual Metadata’, 3–5.

²⁸ Brocks et al., ‘Modeling Context for Digital Preservation’, 215.

²⁹ Lee, ‘A Framework for Contextual Information in Digital Collections’, 104–107, 116–117.

looking for a solution, and the focus on relationships and context makes clear that disconnection could have a negative impact over time.

Different standards and schema for metadata across collection types or institutions can also change the type of information that is captured.³⁰ Bearman makes a key point that a user should not have to know the internal organisation of metadata in order to discover a resource.³¹ This is also reflected by Trant, who suggests that metadata could support connections across museums and their objects, which would free users from "...the boundaries of historical collecting patterns".³² The web has made the idea of linking information an expectation for users, and this expectation extends to their experience of museum collections.³³

The creation of digital surrogates, digital curation, and digital preservation

Digitisation, by definition, refers to the conversion from an analogue format to a digital format.³⁴ Often the term digitisation is used synonymously with digital imaging, and the concept of 'databasing,' or text transcription, is used to refer to the digitisation of written content.³⁵ This can sit alongside digital imaging as a way to provide more access to the digital object. Another key concept is format-shifting. This clarifies that digital media may be stored in collections on an ageing or obsolete physical medium. This is referred to as physical digital media.³⁶ For preservation and access, it will require transferring to a new digital form. This is often called digitisation, and while technically inaccurate, it often follows the same principles and

³⁰ Blackaby and Sandore, 'Building Integrated Museum Information Retrieval Systems', 118.

³¹ David Bearman, 'Possible Contributions of the Reference Model of Metadata Required for Evidence to a Reference Model of Metadata Required for Image Description', *Archives and Museum Informatics* 10, no. 3 (December 1996): 301.

³² J. Trant, 'When All You've Got Is "The Real Thing": Museums and Authenticity in the Networked World', *Archives and Museum Informatics* 12, no. 2 (1 June 1999): 107, 118–119, 121.

³³ Herbert Van de Sompel and Patrick Hochstenbach, 'Reference Linking in a Hybrid Library Environment: Part 1: Frameworks for Linking', *D-Lib Magazine* 5, no. 4 (April 1999): 3.

³⁴ Terras, 'The Rise of Digitization', 3.

³⁵ Randy Singer and Gil Nelson, 'The Impact of Digitization and Digital Data Mobilization on Biodiversity Research and Outreach', (29 August 2018), <https://twitter.com/SiobhanLeachman/status/1034547756189212672>.

³⁶ Moran, 'Born Digital in New Zealand: Report of Survey Results', 15–16.

processes as moving from the analogue to the digital. Digitisation is a useful term, as it encompasses the creation of digital surrogates, or format-shifting of material, for both access and preservation. Issues relating to copyright can affect the ability to do this, however, and library and archive exceptions under the law are not extended to museums and galleries in New Zealand, nor in many international jurisdictions.³⁷ The rise in digital technology is also not well accounted for in the legislation, and this can slow down the process and contribute to format obsolescence and loss.³⁸

Digital preservation and access are important for digital collections long-term, and digital surrogates are often created to address access and preservation concerns for physical collections as well. However, Ciurea and Filip point out that digital imaging does not reduce the desire to see the original.³⁹ This is an area that needs further research, however, as other authors suggest digital imaging to “reduce the pressure on the physical object”.⁴⁰ Many authors agree that the transformation of the physical object into a digital form offers opportunities for access, preservation, and further use.⁴¹ Equally, while digital surrogates cannot preserve everything, they may offer the only way to retain any of the original.⁴² The link to preservation has also been

³⁷ Tim Padfield, ‘Preserving and Accessing Our Cultural Heritage - Issues for Cultural Sector Institutions: Archives, Libraries, Museums and Galleries’, in *Copyright and Cultural Heritage: Preservation and Access to Works in a Digital World*, ed. Estelle Derclaye (Cheltenham, UK; Northampton, MA, USA: Edward Elgar, 2010), 197.

³⁸ Paul Torremans, ‘Archiving Exceptions: Where Are We and Where Do We Need to Go?’ in *Copyright and Cultural Heritage: Preservation and Access to Works in a Digital World*, ed. Estelle Derclaye (Cheltenham, UK; Northampton, MA, USA: Edward Elgar, 2010), 113–114.

³⁹ Cristian Ciurea and Florin Filip, ‘New Researches on the Role of Virtual Exhibitions in Digitization, Preservation and Valorization of Cultural Heritage’, *Informatica Economica* 20, no. 4 (2016): 30; Gould et al., *IFLA/UNESCO*, 27.

⁴⁰ Paul Conway, ‘Digitizing Preservation’, *Library Journal* 119, no. 2 (1994): 45; Gould et al., *IFLA/UNESCO*, 26; N. Mani, ‘Digitisation: Preservation and Challenges’, *DESIDOC Journal of Library & Information Technology* 29, no. 1 (2009): 70.

⁴¹ Ciurea and Filip, ‘New Researches on the Role of Virtual Exhibitions’, 31; Henning, ‘New Media’, 306, 309; Nicholas Jardine, ‘Reflections on the Preservation of Recent Scientific Heritage in Dispersed University Collections’, *Studies in History and Philosophy of Science* 44, no. 4 (2013): 741; Kate Louise Musker Reynold, ‘The Digital Initiative in Archives: A Study on the Selection of Archives for Digitisation in New Zealand’ (Research Paper MLIS, Victoria University of Wellington, 2006), 19, 27–29; Brian Robinson and Simon Tanner, ‘Higher Education Digitisation Service: access in the future preserving the past -- the UK perspective’, *Bibliothek Forschung und Praxis* 23, no. 3 (1999): 66.

⁴² Reynold, ‘The Digital Initiative in Archives’, 21; Marilyn Deegan and Simon Tanner, *Digital Futures: Strategies for the Information Age* (London: Library Association, 2002), 187.

pointed out as an area of opportunity for digital collections by Bremer-Laamanen, suggesting a better link between preservation and digitisation programmes.⁴³

The push towards a greater uptake of digital collections is often attributed to a lack of physical storage space.⁴⁴ Digital collections generally take up less physical space and so can be seen as a solution for storage concerns. However, many authors advise that this may lead to similar issues in the future. Instead, they suggest better collection curation, seen in decisions made during acquisition and in reducing the collection through deaccessioning. While particularly important for physical collections, the literature on digital collections also strongly advocates for this approach.⁴⁵ Knell has a focus on curation in general, discussing connoisseurship and rejecting the notion of perpetuity.⁴⁶ In comparison, Ray discusses digital curation in particular, but also looks at the notion of preservation and longevity, reflecting Knell's sentiment that the cost of keeping collections, whether physical or digital, should not be underestimated.⁴⁷ Equally, they posit that preservation should be thought about from the beginning.⁴⁸

Much of the literature looks at two stages of a digitisation or digital imaging process. First the process itself, and then on digital content preservation. Both of these discussions focus on resource and standards, often looking at the technical considerations and then how these fit into the practicalities of the GLAM environment. Digital preservation is discussed as part of an active process. Conway, speaking from a library perspective, makes this point clear; "in the language of digital imaging and electronic resources, preservation is a verb".⁴⁹ Certainly, the

⁴³ Majlis Bremer-Laamanen, 'Digitisation for Access to Preserved Documents', *Liber Quarterly: The Journal of European Research Libraries* 13, no. 2 (2003), 138.

⁴⁴ Henning, 'New Media', 308; Robinson and Tanner, 'Higher Education Digitisation Service: access in the future preserving the past -- the UK perspective', 66.

⁴⁵ Paul Conway, 'Rationale for Digitization and Preservation', in *Museums in a Digital Age*, ed. Ross Parry (Florence: Routledge, 2010), 373; Sarah Higgins, 'Digital Curation: The Emergence of a New Discipline', *International Journal of Digital Curation* 6, no. 2 (2011): 80; Bremer-Laamanen, 'Digitisation for Access to Preserved Documents', 137.

⁴⁶ Simon J. Knell, *Museums and the Future of Collecting* (Florence: Routledge, 2004), 15–17.

⁴⁷ Knell, *Museums and the Future of Collecting*, 11, 37.

⁴⁸ Ray, 'The Rise of Digital Curation and Cyberinfrastructure', 607.

⁴⁹ Conway, 'Digitizing Preservation', 45.

need for digital preservation, as well as discussions of the technology required for this has been covered in the literature well. The focus on digital preservation also includes clear direction and research into quality standards for imaging work that takes place, in order to ensure long term preservation of the digital file. Much of this has come out of libraries and archives literature, particularly in defining minimum imaging standards. This also extends beyond imaging and includes audio and other formats and file types. Conway has also argued that “access to digital surrogates generates the need for preservation”, as over time these are used as the primary point of inquiry.⁵⁰

As Mani points out, there is a need to preserve accessibility.⁵¹ Sabharwal notes that digital resources have already become less useful because their interactivity was not preserved, in this case ‘dead’ hyperlinks causing non-functionality and a break-down in the relationship between information.⁵² Besser noted in 2000 that preserving the ‘informational content’ requires different thinking than with physical object preservation, as this “may be completely disembodied from any physical artefact”.⁵³

Digital authenticity

A concern around the authenticity of digital objects is noted by a number of authors.⁵⁴ Fleischhauer and Knell both point to the physical objects providing greater credibility for digital surrogates.⁵⁵ By retaining the connection to the physical object, digital surrogates can prove their own authenticity. Knell specifically points out that there is often a “required physicality” when judging authenticity and evidence, which means that for digitisation and imaging to occur without retention of the original

⁵⁰ Conway, ‘Digital Transformations and the Archival Nature of Surrogates’, 55.

⁵¹ Mani, ‘Digitisation: Preservation and Challenges’, 71.

⁵² Arjun Sabharwal, ‘Information Architecture and Hypertextuality: Concerns for Digital Curation’, in *Digital Curation in the Digital Humanities*, ed. Arjun Sabharwal (Chandos Publishing, 2015), 72.

⁵³ Besser, ‘Digital Longevity’, 164–76.

⁵⁴ Henning, ‘New Media’, 307.

⁵⁵ Carl Fleischhauer, ‘Electronic Information and Digitization: Preservation and Security Challenges’, in *The Strategic Stewardship of Cultural Resources: To Preserve and Protect*, ed. Andrea T. Merrill (Florence: Routledge, 2003), 145; Knell, *Museums and the Future of Collecting*, 4; Simon J. Knell, ‘The Shape of Things to Come: museums in the technological landscape’, in *Museums in a Digital Age*, ed. Ross Parry (Florence: Routledge, 2010), 443.

physical object, there needs to be a core change in belief.⁵⁶ Helfrich also points out that a direct comparison to other things is how authenticity is measured, regardless of format.⁵⁷ This is also related to the literature on digital preservation where distrust of digital longevity privileges the physical object.⁵⁸ Conway discusses how to define quality in digitisation, especially on a large scale, and that higher-quality could then mean that digital surrogates could be used in collection management decision making, serving as replacements for physical volumes.⁵⁹ This supports the idea that the comparison to the physical belies quality and therefore authenticity of the digital surrogate. For the preservation of born-digital elements, the idea of authentic experience is discussed.⁶⁰ Anderson provides the examples that a screenshot, a static image, only captures a part of the content of that page, as it does not capture the interactivity of the website.⁶¹

Parry also discusses the complexities of digital authenticity, which initially came out of dealing with art.⁶² Parry suggests adding the term e-tangible, to tangible and intangible, in order to better encompass “a broader definition of objects that allows them to be in a state of motion that occupies different media”.⁶³ This is a useful idea to give digital objects a type of materiality, whilst still differentiating them from tangible physical objects. As Findlay notes, “there is a relationship between an image and the physical object which it records, but heritage is not tangible in digital form”.⁶⁴ Parry notes that authenticity can equally be an issue with physical objects,

⁵⁶ Knell, *Museums and the Future of Collecting*, 4.

⁵⁷ Helfrich, ‘Questions of Authenticity’, 27.

⁵⁸ Mike Jones, ‘Physical Office, Digital Outhouse’, *Context Junky* (blog), 28 July 2015, accessed 28 November 2019, <http://www.mikejonesonline.com/contextjunky/2015/07/28/physical-office-digital-outhouse/>.

⁵⁹ Paul Conway, ‘Archival Quality and Long-Term Preservation: A Research Framework for Validating the Usefulness of Digital Surrogates’, *Archival Science* 11, no. 3 (1 November 2011): 302.

⁶⁰ Jenny Mitcham, ‘Emulation for Preservation - Is It for Me?’ *Digital Archiving at the University of York* (blog), 23 June 2017, accessed 11 February 2020, http://digital-archiving.blogspot.com/2017/06/emulation-for-preservation-is-it-for-me_27.html.

⁶¹ Maxwell L. Anderson, ‘Electronic Information and Digitization: Preservation and Security Challenges’, in *The Strategic Stewardship of Cultural Resources: To Preserve and Protect*, ed. Andrea T. Merrill (Florence, United Kingdom: Routledge, 2003), 164.

⁶² Ross Parry, *Recoding the Museum: Digital Heritage and the Technologies of Change* (London, United Kingdom: Routledge, 2007), 61, 64.

⁶³ Parry, *Recoding the Museum*, 68.

⁶⁴ Peter Findlay, ‘The UK Medical Heritage Library and the Relationship Between Print and the Digital’, *Journal of Victorian Culture* 23, no. 2 (27 April 2018): 236.

particularly in exhibition and display.⁶⁵ This is reflected in a Washington Post article which critiques how clear museums are about the authenticity of physical items on display.⁶⁶ It sets the authenticity of artefact against the authenticity of story and suggests items in the digital space are equal to replicas in a physical space.⁶⁷

Resource

Also pointed out, as an area of disconnection, is the area of resource. However, missing from the literature is an exploration into resource allocation, other than an awareness that resource into digital collections has increased. This may be due to the difficult nature of undertaking this research, however, this prevents further work into the longer-term impact on collection care. In the New Zealand context, the 2017 survey into born-digital collections includes concerns over resource for both physical and digital collections, suggesting that a static resource is now being stretched further.⁶⁸ The literature does, however, look at staff resource and concerns around the future impact treating digital and physical collections in different ways may have on collection care and management. Wellington and Oliver note the siloing of professionals into different areas, and how this contributes to different ways of working for each type of collection, physical and digital, eventually leading to these being two separate collections, disconnected entirely.⁶⁹ My own experience in documentary heritage supports this idea of staff working with each collection coming from very different backgrounds. Equally, however, I have seen collections teams grappling with adding the care of digital collections to their workload, as well as having to upskill to meet this demand.⁷⁰ This is borne out in the New Zealand literature with Goss discussing staff digital capability in archives, and in Orr's research which looks at the challenges that library staff have in having to upskill

⁶⁵ Parry, *Recoding the Museum*, 76.

⁶⁶ Menachem Wecker, 'Are Museums Being Clear Enough with the Public about What's Real and What's Fake?' *Washington Post*, 27 February 2019, accessed 3 April 2019, <https://www.washingtonpost.com/news/magazine/wp/2019/02/27/feature/are-museums-being-clear-enough-with-the-public-about-whats-real-and-whats-fake/>.

⁶⁷ Wecker, 'Are Museums Being Clear Enough with the Public about What's Real and What's Fake?'

⁶⁸ Moran, 'Born Digital in New Zealand: Report of Survey Results', 19.

⁶⁹ Shannon Wellington and Gillian Oliver, 'Reviewing the Digital Heritage Landscape: The Intersection of Digital Media and Museum Practice', in *The International Handbooks of Museum Studies*, eds. Sharon Macdonald and Helen Rees Leahy (Oxford, UK: John Wiley & Sons, Ltd, 2013), 586.

⁷⁰ Ross Parry, *Museums in a Digital Age* (Florence: Routledge, 2010), 1.

themselves, including the isolation they can sometimes feel.⁷¹ The 2017 survey on the state of born-digital collecting also noted that staffing and staff expertise is a key challenge for the future.⁷²

A number of informal case studies on staff allocation also point out that staff are required to upskill themselves in order to drive the uptake of digitisation, or other digital projects, within their organisations, and in dealing with digital preservation.⁷³ Also seen are discussions of separate digital teams, either general information technology staff brought on to support collections or digitally specialised collections staff.⁷⁴ The impact of staff on collections is also mentioned with regards to institutional knowledge. Jones notes that collections knowledge is often held by staff and can be lost with them.⁷⁵ This affects the existing relationships collections have to one another, with Jones pointing to an example where objects in a collection have become disconnected from the field diaries associated with them, and therefore their context, as well as other objects which were collected from the same area.⁷⁶ This alludes to possible disconnection that could occur with digital collections, and that current relationships may be being maintained by institutional knowledge.

⁷¹ Suzanne Marie Goss, 'Developing Digital Capability: What Archivists Can Learn from the GLAM Sector' (Victoria University of Wellington, 2017), 45–47; Renee Orr, 'Digitisation and Workplace Learning: An Exploratory Study' (Victoria University of Wellington, 2006), 39–43.

⁷² Moran, 'Born Digital in New Zealand: Report of Survey Results', 22.

⁷³ Adam Moriarty, 'A Crisis of Capacity: How Can Museums Use Machine Learning, the Gig Economy and the Power of the Crowd to Tackle Our Backlogs – MW19 | Boston' (MW19, Boston, 2019), accessed 20 May 2019, <https://mw19.mwconf.org/paper/a-crisis-of-capacity-how-can-museums-use-machine-learning-the-gig-economy-and-the-power-of-the-crowd-to-tackle-our-backlogs/>; Jeffrey M. Field, 'Building a National Preservation Program: National Endowment for the Humanities', in *The Strategic Stewardship of Cultural Resources: To Preserve and Protect*, ed. Andrea T. Merrill (Florence, United Kingdom: Routledge, 2003), 65; Elvia Arroyo-Ramírez et al., "'Tell Us about Your Digital Archives Workstation": A Survey and Case Study', *Journal of Contemporary Archival Studies* 5, no. Article 16 (2018).

⁷⁴ Matthew Burgess, 'Digital preservation at the point of acquisition: Collecting born-digital photographs', State Library of NSW, 11 February 2019, accessed 4 April 2019, <https://www.sl.nsw.gov.au/blogs/collecting-born-digital-photographs>; Mike Jones, 'On Digital Archives', *Context Junky* (blog), 21 November 2012, accessed 4 December 2017, <http://www.mikejonesonline.com/contextjunky/2012/11/21/on-digital-archives/>; Jessica Moran, 'The Days of Our (Digital) Lives', *Blog | National Library of New Zealand* (blog), 3 October 2018, accessed 24 January 2019, <https://natlib.govt.nz/blog/posts/the-days-of-our-digital-lives>; Parry, *Recoding the Museum*, 43; Gabriela Redwine et al., 'Born Digital: Guidance for Donors, Dealers, and Archival Repositories' (Washington, DC: Council on Library and Information Resources, October 2013), <https://www.clir.org/pubs/reports/pub159/>, 13.

⁷⁵ Jones, 'Documenting Artefacts and Archives in the Relational Museum', 247.

⁷⁶ Jones, 'Documenting Artefacts and Archives in the Relational Museum', 165–166.

Conclusion

While the importance of the relationship between digital surrogates and physical objects is touched on across the literature, little is written about how representing or maintaining this could be achieved, or how it is currently being done. Metadata information across domains alludes to this idea, but this is disconnected from literature which talks about the implications for losing this relationship and context. Disconnection between collections is flagged as an area of future interest in the museums literature, due to the possible effect on each collection that this could pose. Little research has been done on the nature of the relationship, and to what extent digital and physical collections are connected, or disconnected, in museum practice. Equally, the implications and outcomes of this have not been researched. While the literature touches on the relationship between these two collections, rarely is this explored in any depth. The journey from the physical object to a digital surrogate is followed, looking at the technical and practical aspects, but appears as a one-way relationship in the museums literature, which does not echo my own practical experiences working with documentary heritage. Further research is needed to explore in what ways these two types of collections are connected and disconnected and the implications of this for collection care, curation, and the longevity of collection objects.

Specialised language definitions

Some of the terms I used for this research are specialised and may have a variety of meanings depending on the context. It may also be that there are different understandings of these terms across different domains. The following common terms are defined in this research as follows:

Digitisation/Digital imaging

Digitisation is the conversion from an analogue format to a digital format.⁷⁷ Digital imaging is a form of digitisation which refers to the process of creating digital images, including digital image processing.⁷⁸

Digital surrogate

Digital surrogates are the outputs of digitisation and digital imaging.⁷⁹ The digital assets that are created are digital representations of the analogue objects.⁸⁰ It covers all digital representations of physical collections items. These are usually digital images but could also include 3D scans or videos of collection objects.⁸¹ The use of the word surrogate does not indicate that the object is intended to be a substitute of the object, rather it acts as a proxy or representation.

Metadata

Metadata is “data that describes and gives information about other data”.⁸² It is described by Duval et al. as “a key part of the information infrastructure necessary to help create order in the chaos of the Web, infusing description, classification, and organization to help create more useful stores of information”.⁸³

⁷⁷ Terras, ‘The Rise of Digitization’, 3.

⁷⁸ “Glossary; Term: Digital Imaging”, Federal Agencies Digital Guidelines Initiative, accessed 29 January 2020, <http://www.digitizationguidelines.gov/term.php?term=digitalimaging>.

⁷⁹ Cyndi Shein and Emily Lapworth, ‘Say Yes to Digital Surrogates: Strengthening the Archival Record in the Postcustodial Era’, *Journal of Western Archives*, 7, no. 1 (2016): 2.

⁸⁰ Andrea Wallace and Ronan Deazley, ‘Digital Surrogates’, *Display At Your Own Risk*, 2016, accessed 5 February 2020. <https://displayatyourownrisk.org/digital-surrogates/>.

⁸¹ Margot Note, *Managing Image Collections*, Chandos Information Professional Series (Oxford; Cambridge; New Delhi: Chandos Publishing, 2011), Chapter 2.

⁸² “Metadata”, OED Online, accessed 29 January 2020. Oxford University Press, <https://www.oed.com/view/Entry/117150?redirectedFrom=metadata#eid37413841>.

⁸³ Duval et al., ‘Metadata Principles and Practicalities’, 15.

In the museum context, it refers to a wide variety of information that sits around and is about, collection items. I mainly focused on metadata that is contained within a collection record which has a variety of metadata fields, as this is the most accessible and used metadata. However, it also includes other types of metadata such as more technical metadata or filenames.

Source materials

This term was created for the purposes of this research. It refers to the object which a digital surrogate was created from. In this research, this object is usually assumed to be a physical collection object, but it could be any format and include outputs of reformatting projects, as well as digitisation or digital imaging. This term allows for talking across domains, using a format neutral term.

Research questions

While some aspects of digital collections are beginning to be addressed in the museums literature, there are still a number of areas that have not been explored. One area is the connection, or lack thereof, between physical and digital collections, in practice. As digital surrogates are created for collection items and presented in online interfaces to the public alongside born-digital items, the distinction between physical collections and digital collections can become blurred. Also, different teams and processes can be independently responsible for each collection. In the New Zealand context, research which explores the current state of these two types of collections, and their relationship to each other, could help support further research into the long-term implications for the care of both types of collection. I designed my research questions to provide a good foundation for understanding the relationship between digital surrogates and the source material and to provide a strong base for further research, particularly in the New Zealand context, into the long-term effects of current digital imaging and digitisation projects and the metadata creation that sits alongside them.

The primary research question of this thesis is: **In what ways is the relationship between physical collection items and their digital surrogates represented using metadata?**

In order to answer the primary research question, I looked to answer a number of secondary questions, including:

What metadata is created and recorded in the creation of digital surrogates that links it back the physical object?

This question gave me clear information around what metadata sits with the digital surrogates to link them back to the source material, as well as an insight into similarities and differences between the digital surrogate's record and the collection object's record. It also highlighted whether this metadata explicitly represents the relationship or if it implies the digitisation process has occurred.

What are the documented policies around the creation of metadata that links physical objects with their digital surrogates?

This question highlighted the documented policies relating to the creation of metadata in general and whether there is an established practice for representing the relationship between a digital surrogate and physical object using metadata. Gaining an understanding of the documented policies also allowed me to identify any areas where practice may differ from these.

In what ways does metadata represent the relationship, in practice?

This question looked at how metadata links are acknowledged or used, and if people who work directly with the material, and are aware of the relationship already, think that the metadata represents the relationship effectively. Part of answering this question involved answering much smaller questions such as, which fields are used, what is and is not shown in the public interface, and are there non-documented ways of representing the relationship? This question also highlighted any differences in their view or approach for staff in different roles in the organisation.

Does the purpose (access/preservation) of the digital surrogate impact or effect the way metadata is created, recorded, or displayed?

This question highlighted whether metadata, which represents the relationship, is impacted or affected by the reasons for digitisation and explores the drivers for digitisation, most commonly access and preservation.

By looking at the way these relationships are managed I invoked the idea of collections management. This covers the care of these collections, including preventative preservation, housing and storage, and the administrative care of these collections, including collection records and documentation. I looked at management as it refers to the organisation and control of collections and collection objects, expecting that the relationship will be managed through staff communication or records and administration policies and systems. This specifically included how these are or are intended to be, sustained over time, both physically and intellectually.

Research design

This research looked to explore how the relationships between digital surrogates and their source material are being represented using metadata under the assumption that if they are not, there are implications for the future context of these collection items. There was an awareness that other ways may be used to represent the relationship and context for these collection items. As noted in the literature review, connections between differently managed physical collections can lead to a breakdown in contextual information that negatively impacts audience and institutional understanding of the objects, as it does not support long term retention of knowledge. If not managed now, or if managed in a non-sustainable way, institutions cannot ensure that the relationships will remain over time and that their physical and intellectual linkages will be preserved. By understanding how a large collecting institution like Te Papa currently uses metadata to represent the relationship between collection objects and digital surrogates, a better understanding of this area may be gained. The future implication for current practice may be better understood as time progresses, and getting a current state of practice at one institution could provide a point of comparison for differing approaches leading to knowledge sharing and best practice standards over time.

This research is important because the relationship between a physical object and its digital surrogate gives the digital surrogate meaning and context. It also cannot fulfil any of its possible purposes, either for access or preservation, if this relationship is broken, even if the digital surrogate is retained as a digital object in its own right. Metadata functions as a clear way to maintain this relationship and offers many ways to achieve this. By exploring what is happening at a large New Zealand museum, like Te Papa, which is likely to be influencing sector practice, this research sought to understand current practice, which supports future research into the long-term implications for sustainability, as well as changes in practice over time.

Case study as a research strategy

Case studies are a useful research strategy for understanding complex relationships and provide depth and insight into a discrete instance.⁸⁴ This approach was chosen as the research strategy to address my research question as it has a primary focus on relationships, looking at not only what is happening but also exploring why those approaches have been taken.⁸⁵ A case study allowed for a more holistic view, as well as multiple methods which led to a better understanding of the research question.⁸⁶ The purpose of the case study was primarily one of description, with some exploration of the issues affecting what is being described.⁸⁷ The focus was on process, rather than outcomes.

The case study did not uncover transferable findings, as organisations in New Zealand, and internationally, are unlikely to have similar features or particulars.⁸⁸ It is part of “an exploratory foundation that helps with the development of theory”, and analysed the situation at my case study site.⁸⁹ This analytic aspect of the findings was then generalised, as it contributes to broader theory. A disadvantage of a case study can be the lack of credibility in generalising any findings, however, for this research, I focused on analytic generalisation, which is an approach that focuses on concepts and theories.⁹⁰ While the specific context cannot be generalised, instead the theories and concepts developed during the data analysis stage of the research have a level of fluid generalisation and are understood as needing corroboration through further research.⁹¹

⁸⁴ Martyn Denscombe, *The Good Research Guide: For Small-Scale Social Research Projects* (England: McGraw-Hill/Open University Press, 2014), 4, 54.

⁸⁵ Denscombe, *The Good Research Guide*, 55.

⁸⁶ Denscombe, *The Good Research Guide*, 56.

⁸⁷ Denscombe, *The Good Research Guide*, 57.

⁸⁸ Denscombe, *The Good Research Guide*, 62–63.

⁸⁹ Denscombe, *The Good Research Guide*, 61.

⁹⁰ Denscombe, *The Good Research Guide*, 64; Brett Smith, ‘Generalizability in Qualitative Research: Misunderstandings, Opportunities and Recommendations for the Sport and Exercise Sciences’, *Qualitative Research in Sport* 10, no. 1 (October 2018), 5.

⁹¹ Denscombe, *The Good Research Guide*, 61; Smith, ‘Generalizability in Qualitative Research: Misunderstandings’, 5.

An ideal approach to this would have been to do two case studies. This would have provided me the opportunity to compare the organisations and their approaches and see if similar or different ideas and themes emerged. This would have given the research a much more in-depth look at the New Zealand context, particularly large, well-resourced institutions. Case studies which compared differently sized organisations with different governance structures would also allow this comparison and have led to greater transferability of my findings. However, this would also be very time and resource intensive, and doing a single site case study still provided the research with the depth it needed to better understand the topic and why themes or ideas emerged, despite not being able to draw certain conclusions about the phenomena. One limitation of a case study is that it does not give a broad overview of the current state, however, other research strategies, did not provide the opportunity to look “in depth at the subtleties and intricacies” of the phenomena I encountered.⁹²

When considering the research strategy others were considered, particularly using a survey to gain a broader understanding across museums in New Zealand. This approach would have provided wide coverage and an understanding of what is happening in different institutions, but with less focus on why they were operating this way and how their individual approaches impacted on the way they used metadata.⁹³ The research question focuses on gaining depth of understanding, as it explores the context for decisions made relating to metadata and relationships, and a case study was chosen to achieve this research aims.⁹⁴ A case study allowed for a deeper understanding within the restricted time-frame of a Master’s. The research questions required a good understanding of relationships and links, which a case study allowed me to achieve, by employing more than one method and allowing me to look in detail at each aspect.

⁹² Denscombe, *The Good Research Guide*, 63.

⁹³ Denscombe, *The Good Research Guide*, 7–8.

⁹⁴ Denscombe, *The Good Research Guide*, 4.

Data collection methods

The two methods of data collection that were used were interviews and documentary research. Observation was also considered to gain a sense of the actual practice of metadata creation, however, interviews provided a better method for understanding and exploring why certain actions were taken. My research questions also looked at a specific aspect of this process and so observation would have been a very time-consuming process, without gaining relevant, quality data.⁹⁵

For the interviews, I used exploratory sampling, rather than representative sampling.⁹⁶ Exploratory sampling does not seek to gain a cross-section of the research population, but rather looks to use specific examples to illuminate the research topic.⁹⁷ This involved snowball sampling, where an initial participant was approached and also asked to nominate others who fit the criteria that were sought after for the research.⁹⁸ This generated a non-probability sample which meant I was able to select for expertise and professional knowledge.⁹⁹ This worked alongside purposive sampling, as the criteria for participants was related to their roles and associated expertise and knowledge which gave quality insight and information on the research topic.¹⁰⁰ Using the sampling method I was able to get information-rich participants who could act as key informants; individuals with the ability to offer deeper insight into their area of expertise.¹⁰¹ This approach meant that their individual perspectives and approaches were more visible in the research, however, they provided more direct and much faster access to rich information.¹⁰² This best suited my smaller-scale and qualitative research, particularly as I was not looking to draw generalisations across a research population, but instead discover more about the research area and generate insights.¹⁰³

⁹⁵ Denscombe, *The Good Research Guide*, 212.

⁹⁶ Denscombe, *The Good Research Guide*, 32.

⁹⁷ Denscombe, *The Good Research Guide*, 33.

⁹⁸ Denscombe, *The Good Research Guide*, 42–43.

⁹⁹ Denscombe, *The Good Research Guide*, 34.

¹⁰⁰ Denscombe, *The Good Research Guide*, 41.

¹⁰¹ MN Marshall, 'The Key Informant Technique', *Family Practice* 13, no. 1 (1996): 92.

¹⁰² Marshall, 'The Key Informant Technique', 92–93. Geoff Payne and Judy Payne, 'Key Informants', in *Key Concepts in Social Research* (London: SAGE Publications, Ltd, 2011), 135, 137.

¹⁰³ Denscombe, *The Good Research Guide*, 32–33.

Interviews have the advantage that they provide a great depth of information and insight into my research area.¹⁰⁴ I also had the flexibility to explore participants' priorities, allowing them to explain their views and what they thought was relevant to the research area, which provided useful insights into unique relationships and processes.¹⁰⁵ Interviews can have the disadvantage of being time-consuming and rely on the interviewee for reliability, possibly reflecting only their ideals, not what is actually occurring.¹⁰⁶ By seeking themes from across the interviews I could see where those ideas were shared by a wider group.¹⁰⁷ Documentary research was also used and provided further evidence for these shared themes.¹⁰⁸ An advantage of using documentary research was that it provided me with relatively accessible data which gave context to the interview responses and helped me to find focus areas.¹⁰⁹ A disadvantage of using documentary research was that it is not necessarily an objective account of the actual practices in the organisation, rather the documents act as a permanent record of policies and approaches which may influence, or be influenced by, actual practice.¹¹⁰ For this reason, the documentary research focused on guidelines for practice across the research topic and were looked at alongside the data gathered through the interviews. This ensured that information and insights gained from the interviews that shared themes with documentary sources could be seen in the context of one another.

¹⁰⁴ Denscombe, *The Good Research Guide*, 184, 201–202.

¹⁰⁵ Denscombe, *The Good Research Guide*, 202.

¹⁰⁶ Denscombe, *The Good Research Guide*, 200–201.

¹⁰⁷ Denscombe, *The Good Research Guide*, 201.

¹⁰⁸ Denscombe, *The Good Research Guide*, 225.

¹⁰⁹ Denscombe, *The Good Research Guide*, 225–226, 228–230.

¹¹⁰ Denscombe, *The Good Research Guide*, 226, 230.

Case study selection

A single site case study was chosen to enable an in depth look at a New Zealand institution, which is likely, due to its large size and presence in the sector, to be influencing broader practice. A single site case study provided a detailed look at the particular organisation's activities, including their digitisation and digital imaging workflows, to understand how the relationship between physical objects and their digital surrogates is represented by metadata. As I am interested in relationships and processes, and how these impact the use of metadata, the case study allowed me to explore this, and any problems or opportunities that arose.¹¹¹ It was also important that the site chosen had a focus on rapid digitisation as these projects can bring to the fore the importance of metadata, digitisation, policies, and practices.¹¹²

In selecting a large institution, I was interested in an organisation which employed a large number of people to work across various aspects of the digitisation process. All of the institutions which were considered have specific teams working on digitisation work, rather than it being primarily an upskilled part of a role which has a different primary focus.¹¹³ This indicated a level of resource, or priority of resource, on digitisation. While smaller institutions may also meet these criteria, the focus was on institutions with a greater presence in the sector, that may be more likely to be influencing practice. For Te Papa, this is primarily seen in the National Services Te Paerangi (NSTP) team which offers practical and strategic programmes to the sector.¹¹⁴ For both Te Papa and Auckland War Memorial Museum, their commitment to presenting their practice at local conferences also supports their presence in the sector.¹¹⁵

¹¹¹ Denscombe, *The Good Research Guide*, 57.

¹¹² Field, 'Building a National Preservation Program', 65; Arroyo-Ramírez et al., 'Tell Us about Your Digital Archives Workstation'.

¹¹³ Goss, 'Developing Digital Capability: What Archivists Can Learn from the GLAM Sector', 45–47; Orr, 'Digitisation and Workplace Learning: An Exploratory Study', 39–43.

¹¹⁴ 'National Services Te Paerangi', Museum of New Zealand Te Papa Tongarewa, Wellington, New Zealand, 9 February 2016, accessed 26 February 2020, <https://www.tepapa.govt.nz/learn/formuseums-and-galleries/national-services-te-paerangi>.

¹¹⁵ 'Past conferences', National Digital Forum, accessed 26 February 2020. <http://www.ndf.org.nz/past-conferences-1>.

The idea of multiple case studies was considered. This may have offered the ability to compare different or similar approaches. Comparing two museums was explored, as was comparing Te Papa with the Alexander Turnbull Library, providing data across domains. This last option was removed as it expanded the scope of the research too widely, requiring a much more in-depth literature review which would have explored each of the museum, library, and archive domains with a much greater depth than a Master's thesis would allow. Elements of these domains have been explored in the literature. A single site case study was ultimately chosen in order to gain more depth of information from a single institution and to act as a snapshot of current practice in a museum space, rather than a more comparative approach which would have shown a more surface level analysis of differing institutions or domains.

I reached out to both Te Papa and Auckland Museum. These institutions had recently, or were currently, undertaking large digital imaging projects involving many teams. Both organisations are increasingly incorporating digital interventions into their practice and are invested in initiatives to support other institutions. Choosing the institution involved conversations, either in person or via email, with a person at the institution who was able to describe the types of people working there and the sort of documentation I would have access too. Both conversations were very positive; and Te Papa was chosen as the case study as they were in the midst of their rapid digitisation project. Moreover, as the national museum, they had the size and resource I was looking for.

The chosen case study served as the starting point for further research, and in the development of theory in the area of digital collections.¹¹⁶ This research provided a good understanding of the current state in a large well-funded New Zealand institution, and provided the rationale for improvements in professional practice, supported by theory.

¹¹⁶ Denscombe, *The Good Research Guide*, 61.

Data collection

Data was collected through one-on-one interviews with staff at the institutions, as well as documentary research, using collection plans and policies, and cataloguing manuals. Data collection took place from October – November 2019. Documentary research occurred across this period with interviews taking place in October.

Interviews

My interviews were conducted at Te Papa and took between 60 and 90 minutes. I used a semi-structured interview style, using an interview guide as the basis for all the interviews (see Appendix one). This meant that questions were covered in a similar order but there was flexibility around the questions asked, and the order of the questions based on the participant's answers. Not all of these questions were posed to all of my interviewees and other questions were also asked to better understand an answer, or prompt for further depth. The interviews were recorded, with limited field notes written to supplement the recording. The recordings were transcribed and then summarised before being analysed.

I spoke to four people across different teams and roles. By talking to people in different teams I was able to get a range of opinions and points of view from different museum professionals. This helped provide a number of perspectives, and so provided a good amount of depth of opinion across more of the institution than just the team that deals with digital imaging. I spoke to them as key informants, who were able to speak to the research topic from their particular perspective and area of expertise.¹¹⁷

My aim was to speak to someone in the digital imaging team, who is involved in the creation of digital surrogates; a Collection Manager, who is involved in creating and recording the metadata which represents the relationship; a Curator, who is part of making decisions about what is digitally imaged; and possibly someone from the

¹¹⁷ Marshall, 'The Key Informant Technique', 92.

web team, who is involved in determining which metadata gets displayed on the publicly accessible online collection. In the end, this last role was instead filled by the Collections Information System Manager who was able to provide this information and insight, as well as further information on the collection management system itself.

Because I was interested in speaking to people who do have some knowledge of digital imaging in the museum environment, I used snowball sampling as part of the recruitment process.¹¹⁸ I initially reached out to the Head of Collection Access who put me in touch with my first interviewee, Fiona Moorhead, Collections Information System Manager. She was well placed to talk about the collection management system, as well as digital workflows and Collections Online. I indicated to her the types of roles I was also looking to interview, and she was able to suggest people in those roles who I could speak to, as well as have an initial discussion with them before putting me in touch. My other three interviews occurred because of this process. This aligned with my research plan, as I wanted to get information rich participants who could act as key informants.

Documentary research

I also collected data from documentary sources. These sources included cataloguing manuals, guides to the collection management system (CMS) structure, and digitisation and collection priority documentation. These data sources were important for answering ‘what are the documented policies around the creation of metadata that links physical objects with their digital surrogates?’. They also provided a point of comparison to information from the interviews about actual practice.

I used official documentary sources, which gave me a better idea of the overall focus of the organisation and served as a useful comparison against data gathered during interviews. Table 1 lists the documentary sources which were consulted. This

¹¹⁸ Denscombe, *The Good Research Guide*, 43.

includes policy documents and plans, particularly those specific to collections and digital interventions in the museum, as well as metadata standards. These may be written for a more specialised audience, particularly staff, and provide good data on how collections are actually looked after, resourced, and connected.

I had also intended to complete documentary research where I compared a small number of internal collection records with the public online collection information. In the end, this only occurred in an informal way, as my expectation that specific fields may be used for metadata related to my research questions was ultimately incorrect. I was still able to make these comparisons, but this was less relevant to the research than initially anticipated.

Table 1: Documentary sources

Adrian Kingston. 'KE EMu Media Assets Manual, General Staff Version'. Museum of New Zealand Te Papa Tongarewa, 27 June 2006.
Adrian Kingston and Carol Stevenson. 'Core Cataloguing Guidelines – Humanities Collections'. Museum of New Zealand Te Papa Tongarewa, Version 1.2, July 2015.
Fiona Moorhead, 'EMu Media assets at Te Papa: an overview of the current state, Museum of New Zealand Te Papa Tongarewa, March 2019.
'Draft Priority list June 2019 [ACDP/Photography Collection],' Museum of New Zealand Te Papa Tongarewa, June 2019.
'Photography Collection Plan – identifying tasks, volume of work required, and priorities,' Museum of New Zealand Te Papa Tongarewa, provided October 2019.
'Te Papa Media Assets in EMu: Security and Repositories,' Museum of New Zealand Te Papa Tongarewa, Version 2.9, September 2019.

Conducting interviews also required ethical consideration, including submitting an application to the Human Ethics Committee for approval before undertaking data collection. Informed consent was sought from each of my participants. Due to the

nature of a single site case study, and the selection of participants from particular professional areas, I was unable to offer confidentiality. To account for this, I provided participants with summaries of their interviews for them to look over. This ensured they were aware of what would be attributed to them in the research, including the type of context these comments may be made in. In this process, they were able to remove or clarify points they made in the interview. This may have resulted in changes to the data collected during the interview, including, for example, the removal of candid comments about the process or changes to their responses to interview questions based on information discussed later in the interview. This process provided participants with the ability to control the information they provided, which was important due to my inability to anonymise them in my research, and gave an initial check on the accuracy of data.¹¹⁹ I also provided them the option to be named in the research or to be referred to only by their role. All of my interviewees were happy to be referred to by name.

Data analysis

When analysing the data I used thematic analysis. Thematic analysis is a type of narrative analysis of qualitative data which is based on a coding method.¹²⁰ It is a “method for ordering and synthesising data”, with central and subthemes.¹²¹ ‘Themes’ are usually considered “a common-sense way to refer to patterns in the data that reveal something of interest regarding the research topic at hand”.¹²² King and Horrocks define themes as “recurrent and distinctive features of participants’ accounts, characterising particular perceptions and/or experiences, which the researcher sees as relevant to the research question”.¹²³ At a basic level, a theme can be defined as “a unit of meaning”.¹²⁴ In thematic analysis, data is coded by the researcher to represent themes.¹²⁵ This is done through reading and rereading

¹¹⁹ Denscombe, *The Good Research Guide*, 200–201.

¹²⁰ Alan Bryman, *Social Research Methods*, Fifth edition (Oxford; New York: Oxford University Press, 2016), 553–554, 700.

¹²¹ Bryman, *Social Research Methods*, 554.

¹²² Nigel King and Christine Horrocks, *Interviews in Qualitative Research* (Los Angeles: SAGE, 2010), 149.

¹²³ King and Horrocks, *Interviews in Qualitative Research*, 150.

¹²⁴ Greg Guest, Kathleen MacQueen, and Emily Namey, *Applied Thematic Analysis* (California: SAGE Publications, Inc., 2012), Chapter 3.

¹²⁵ Guest, MacQueen, and Namey, *Applied Thematic Analysis*, Chapter 1.

transcripts and field notes.¹²⁶ These codes have defined boundaries for the theme. It is a flexible approach which allows the researcher to define the core themes in the data.¹²⁷

The first stage of coding was to use descriptive codes for the interview summaries which had been accepted by the interviewees. These codes described what each interviewee was talking about. This was repeated to incorporate more interpretive analysis of the data (See Appendix two, ‘Descriptive codes’). The coding process was repeated for the internal documentation I had acquired, using these same codes, to allow the documents to support or contrast the interviews. The codes were then grouped together into larger more interpretive themes (See Appendix two, ‘Themes’). These themes were mainly drawn from the information coded in the interviews, with supplementary information from the documentary sources. During this process, further documents were also sought from the institution as my awareness of themes in the interview data grew. These were coded and added to the data. The four themes formed the basis of my three findings chapters. This research serves as a starting point for understanding how metadata is used to represent the relationship between a physical object and its digital surrogate in New Zealand.

Thesis structure

The first chapter of this thesis introduces my case study site, providing a background to the organisation and their projects and systems related to the research, and introduces my interviewees. The following chapters explore the three key findings of the research. Chapter two looks at the different worldviews, perspectives, and language encountered in the research, including a comparison between domains. Chapter three explores the system architecture and how it represents the relationship between physical objects and their digital surrogates. How this relationship is represented using metadata is discussed in chapter four which looks to directly answer the main research question.

¹²⁶ Bryman, *Social Research Methods*, 554.

¹²⁷ Bryman, *Social Research Methods*, 700.

Chapter one: Case study - Museum of New Zealand, Te Papa Tongarewa

Introduction to Te Papa

The history of Te Papa as an institution begins in 1865 with the establishment of the Colonial Museum in Wellington.¹²⁸ Modelled on other Victorian-era museums its collections were primarily natural history based.¹²⁹ In 1907 it was renamed the Dominion Museum to reflect New Zealand's Dominion status.¹³⁰ The collections continued to grow and expanded in scope, and the National Art Gallery and Dominion Museum Act was passed in 1930, uniting the organisations under a board of trustees.¹³¹ Buckle Street in the suburb of Mount Cook was chosen as the site for a new building for these entities, alongside a National War Memorial.¹³² Opened in 1936, it began to lose popular appeal, partially due to its location on the city fringe.¹³³ After the building was used by the air force during World War II, it reopened in 1949 with a more inward focus on collections and research.¹³⁴ At this point, both the art and museum collections had grown beyond the capacity of the site and discussions were had about other possible sites for relocation.¹³⁵ Overseas museological trends in the post-war period were beginning to be integrated into New Zealand institutions, who began looking “at the country's own cultural heritage and national identity”.¹³⁶ In 1980 the *Te Maori* exhibition initiated “sweeping changes in museums”, primarily around the inclusion of Māori staff and language.¹³⁷ This, along with social and political progressiveness, formed a dynamic period that the development of Te Papa came out of.¹³⁸

¹²⁸ Conal McCarthy, *Te Papa: Reinventing New Zealand's National Museum, 1998-2018* (Wellington, New Zealand: Te Papa Press, 2018), 29

¹²⁹ McCarthy, *Te Papa*, 30.

¹³⁰ Shannon Wellington, 'Building GLAMour: Converging Practice between Gallery, Library, Archive and Museum Entities in New Zealand Memory Institutions' (PhD, Wellington, New Zealand, Victoria University of Wellington, 2013), 93.

¹³¹ Wellington, 'Building GLAMour', 94.

¹³² Wellington, 'Building GLAMour', 95–96.

¹³³ McCarthy, *Te Papa*, 30, 32.

¹³⁴ McCarthy, *Te Papa*, 32.

¹³⁵ Wellington, 'Building GLAMour', 98.

¹³⁶ McCarthy, *Te Papa*, 32.

¹³⁷ McCarthy, *Te Papa*, 32.

¹³⁸ McCarthy, *Te Papa*, 34–35.

The Museum of New Zealand Te Papa Tongarewa Act 1992 united the National Gallery and the National Museum (renamed in 1972) entities into a single integrated collection and structure.¹³⁹ Approval and funding to house a new institution was received in 1986, and this was opened in 1998 on the current Wellington waterfront site.¹⁴⁰ Changes in society, as well as changes in museums and a new direction for museology, led to a museum which had three core concepts around social inclusion, narrative storytelling, and collaboration with indigenous people at its heart.¹⁴¹ Te Papa’s vision, “to change hearts, minds, and lives”, is one which gives them a wide scope to explore several different stories, presentation methods, and collecting practices.¹⁴² Their stated goals, as they relate to collections, focus on their commitment to those collections, as well as indicating the use of them to support understanding and engagement, particularly in commitment to Te Tiriti o Waitangi.¹⁴³

Early exhibitions innovatively incorporated digital elements and this digital inclusion has continued across the museum.¹⁴⁴ After multiple CEOs, restructures, and general changes in the focus, particularly over the last 10 years, the role of digital has changed across both front- and back-of-house functions. McCarthy, in his work on the first 20 years of Te Papa, notes that a theme that emerges is the importance of digital technology across the museum. He says that “understanding the virtual, alongside the physical, as merely another way of managing artefacts, telling stories and engaging with communities means that museum professionals can add it to their array of tools, without abandoning the analogue, or any other method that is useful,

¹³⁹ McCarthy, *Te Papa*, 29; Wellington, ‘Building GLAMour’, 100.

¹⁴⁰ McCarthy, *Te Papa*, 29; Wellington, ‘Building GLAMour’, 90.

¹⁴¹ McCarthy, *Te Papa*, 40.

¹⁴² ‘Te Papa’s Vision and Future’, Museum of New Zealand Te Papa Tongarewa, Wellington, New Zealand, 10 February 2016, accessed 9 December 2019, <https://www.tepapa.govt.nz/about/what-we-do/te-papas-vision-and-future>.

¹⁴³ “Our goals: New Zealanders are engaged in cultural and contemporary issues through participation in Te Papa events, outreach, exhibitions, and activities; Collections are developed and preserved for present and future users; Iwi and communities are increasingly engaged with their histories, traditions, taonga treasures and collections, in partnership with Te Papa; Visitors are enabled to better understand Aotearoa New Zealand’s heritage, arts, sciences, and culture through Te Papa’s collections, knowledge, and research”. ‘Te Papa’s Vision and Future’.

¹⁴⁴ McCarthy, *Te Papa*, 86–87, 163–164, 177, 181.

however unfashionable”.¹⁴⁵ This space is where this research sits, across the digital and the analogue, and grounded in collection management.

Digitisation and collection management at Te Papa

Te Papa is currently undertaking a rapid digitisation project, the Accelerated Collections Digitisation Project (ACDP), which builds upon its previous digitisation work as part of business as usual processes. The outputs of this work are incorporated into their internal collection management system, EMu, and the public collections website, Collections Online. The digitisation project and the systems that capture and display the collections are part of the digital foundation of collection management at Te Papa.

The Accelerated Collections Digitisation Project

The ACDP was launched in August of 2017 and is a rapid digitisation project, focusing on international best practice standards and making the collections more accessible.¹⁴⁶ The project involves high-resolution imaging, specialised roles, new imaging equipment, as well as work on clearing rights. Little is written publicly about the project, other than a Te Papa Blog to celebrate the 10,000th image, from early 2018, outlining the project and highlighting a range of the images created. It notes that “alongside photographing the collection items in high-resolution, we are also clearing rights and adding contextual information to the records like names, locations, and subjects to make the items more discoverable online”.¹⁴⁷

EMu, Te Papa’s collection management system (CMS)

Te Papa’s current CMS, KE EMu, was implemented in June 2005. A CMS is a program which supports the management of the collections by archiving and cataloguing objects in the collection, and is used to track related information. The

¹⁴⁵ McCarthy, *Te Papa*, 181.

¹⁴⁶ Gareth Watkins, ‘Rapidly Digitising 10,000 Collection Items’, Te Papa’s Blog, 26 February 2018, accessed 22 November 2019, <https://blog.tepapa.govt.nz/2018/02/27/a-digital-birthday-treat/>.

¹⁴⁷ Watkins, ‘Rapidly Digitising 10,000 Collection Items’.

current system includes Media Assets functionality developed by Te Papa and KE Software for Te Papa's version of EMu.¹⁴⁸ Media assets are media related to a collection object, such as a physical negative or photographic print, as well as digital media, including digital images. Prior to using this CMS, they used one called 'Te Kahui' which could not manage digital media. "A thumbnail was linked to a collection object, but the images were stored separately (full-sized images scattered through 650 CDs and thumbnails and screen sized images stored on a server) and there was minimal, if any, information relating to the actual digital files".¹⁴⁹ The media assets functionality that was developed ensured a higher quality of preservation for the digital media and better linked it to the rest of the collection.

The EMu catalogue is divided into three parts; Collection, Media, and Accessories. "The Collection section of the catalogue contains all Te Papa's information about collection objects (see Figure 1). The Media section contains information relating to Te Papa's media assets, i.e. the "information objects" Te Papa (or its predecessors) created or creates as part of its business, that relate to the collections or research (see Figure 2). The Accessories section of the catalogue records information about reusable objects that Te Papa needs to manage in relation to its collections or media, for example, scanners, cameras, mannequins, crates, frames. The Multimedia module is used simply to ingest the file into the system, and as part of that process, read technical attributes into the media asset record (see Figure 3)".¹⁵⁰

¹⁴⁸ Adrian Kingston, 'KE EMu Media Assets Manual, General Staff Version' (Museum of New Zealand Te Papa Tongarewa, 27 June 2006), 2.

¹⁴⁹ Kingston, 'KE EMu Media Assets Manual, General Staff Version', 2.

¹⁵⁰ Kingston, 'KE EMu Media Assets Manual, General Staff Version', 2.

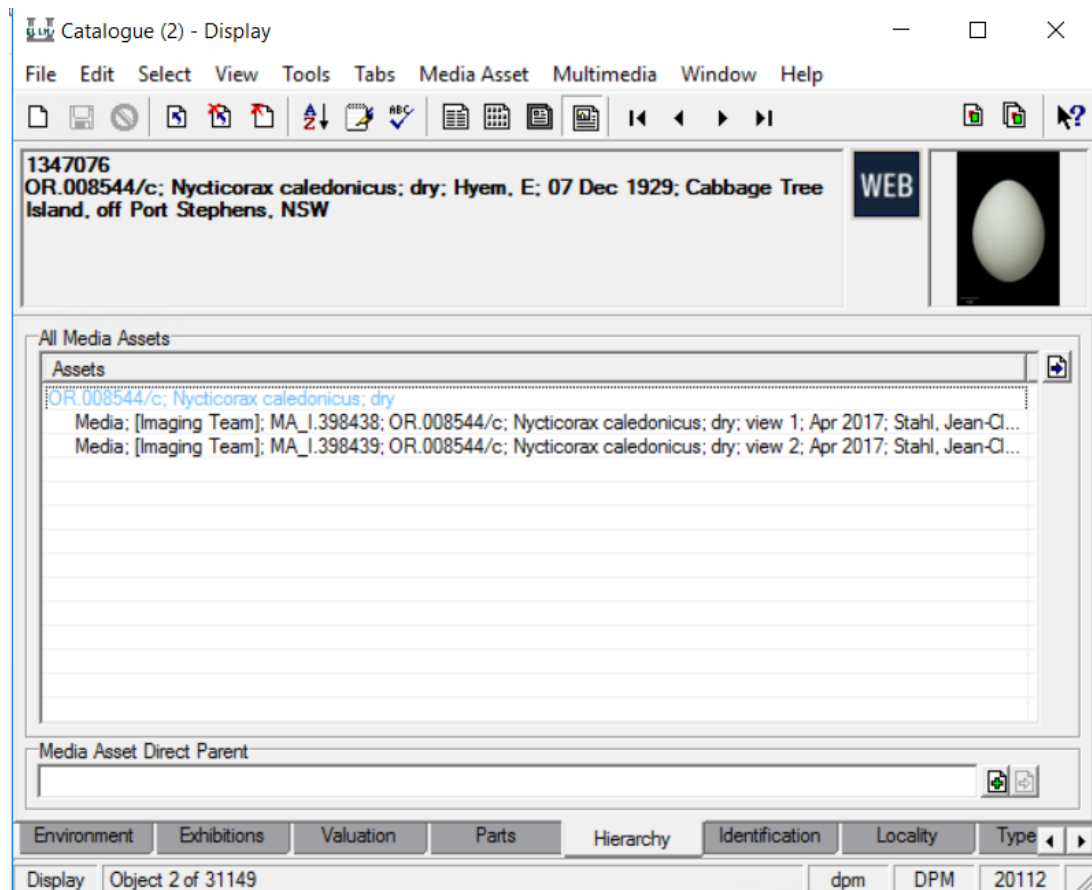


Figure 1: Image of catalogue record in EMu (IRN 1347076). Showing the associated media assets. Taken from ‘EMu Media assets at Te Papa: an overview of the current state, Museum of New Zealand Te Papa Tongarewa, March 2019, 7.

Features collection item: Nankeen Night Heron, *Nycticorax caledonicus*, collected 7 December 1929, Cabbage Tree Island, off Port Stephens, NSW, Australia. CC BY-NC-ND 4.0. Te Papa

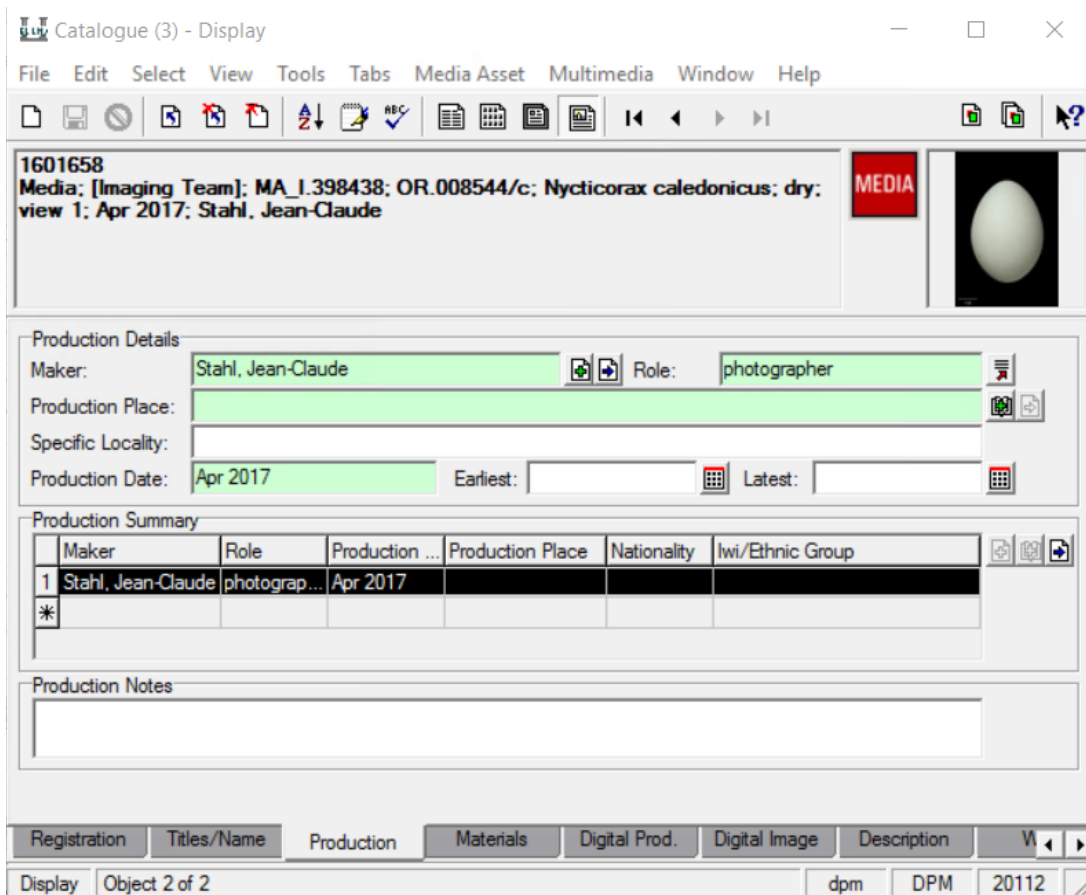


Figure 2: Image of media asset record in EMu (IRN 1601658). Taken from 'EMu Media assets at Te Papa: an overview of the current state, Museum of New Zealand Te Papa Tongarewa, March 2019, 7.

Features collection item: Nankeen Night Heron, *Nycticorax caledonicus*, collected 7 December 1929, Cabbage Tree Island, off Port Stephens, NSW, Australia. CC BY-NC-ND 4.0. Te Papa (OR.008544/c).

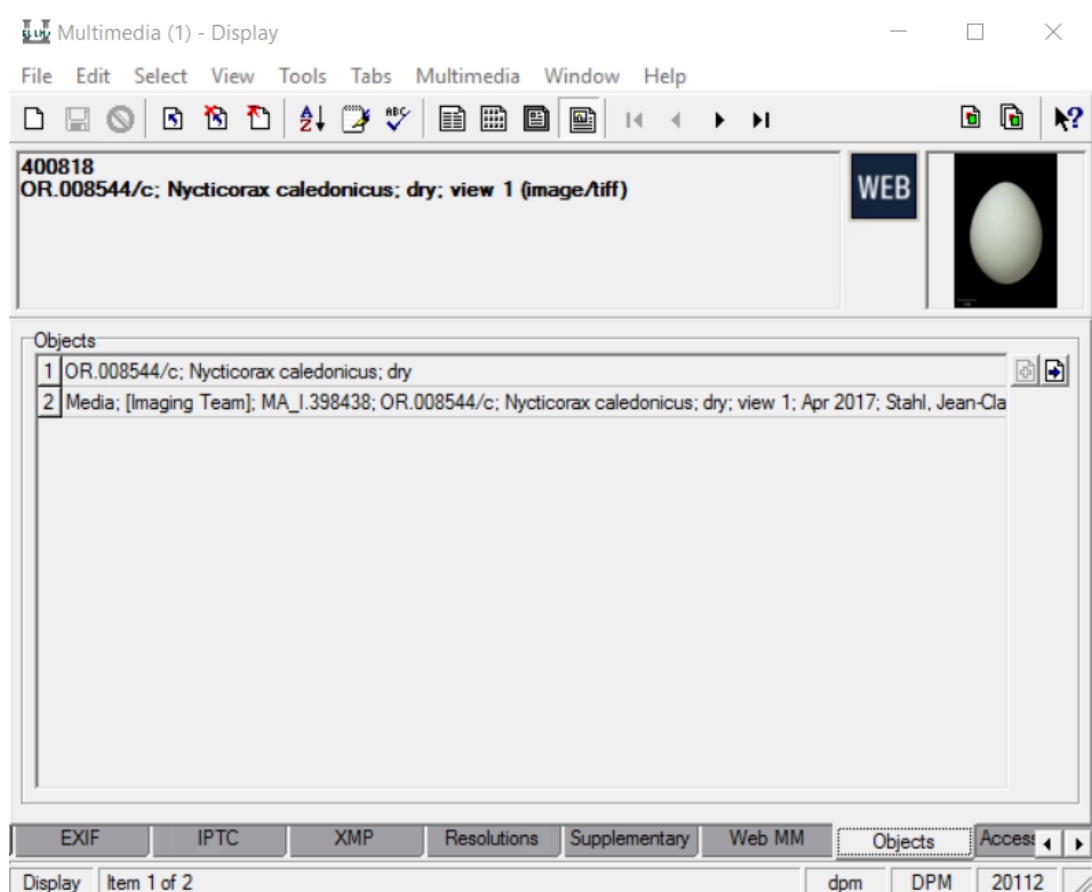


Figure 3: Image of multimedia record in EMu (IRN 400818). Taken from ‘EMu Media assets at Te Papa: an overview of the current state, Museum of New Zealand Te Papa Tongarewa, March 2019, 7. Features collection item: Nankeen Night Heron, *Nycticorax caledonicus*, collected 7 December 1929, Cabbage Tree Island, off Port Stephens, NSW, Australia. CC BY-NC-ND 4.0. Te Papa (OR.008544/c).

Collections Online

Collections Online (<https://collections.tepapa.govt.nz/>) is Te Papa’s online website which provides public access to their collection information including images. It also contains further information to provide context to these items “on related people, places, topics, species, and research from Te Papa”.¹⁵¹ It was launched in September 2017 and completely replaced their previous collections site which was built in

¹⁵¹ ‘About Collections Online’, Museum of New Zealand Te Papa Tongarewa, Wellington, New Zealand, 16 August 2017, accessed 7 January 2020, <https://www.tepapa.govt.nz/discover-collections/collections-online/about-collections-online>.

2009.¹⁵² Collections Online was designed to improve access through searching and browsing of collections and had improved features such as image zoom, support for macrons, and search filters.¹⁵³

Figure 4 shows a search on Collections Online which indicates the types of search filters available on the left, as well as how items are displayed. This is page two of the search and shows a variety of objects created by Rita Angus or featuring her work, as well as links to records about the person and a related organisation.

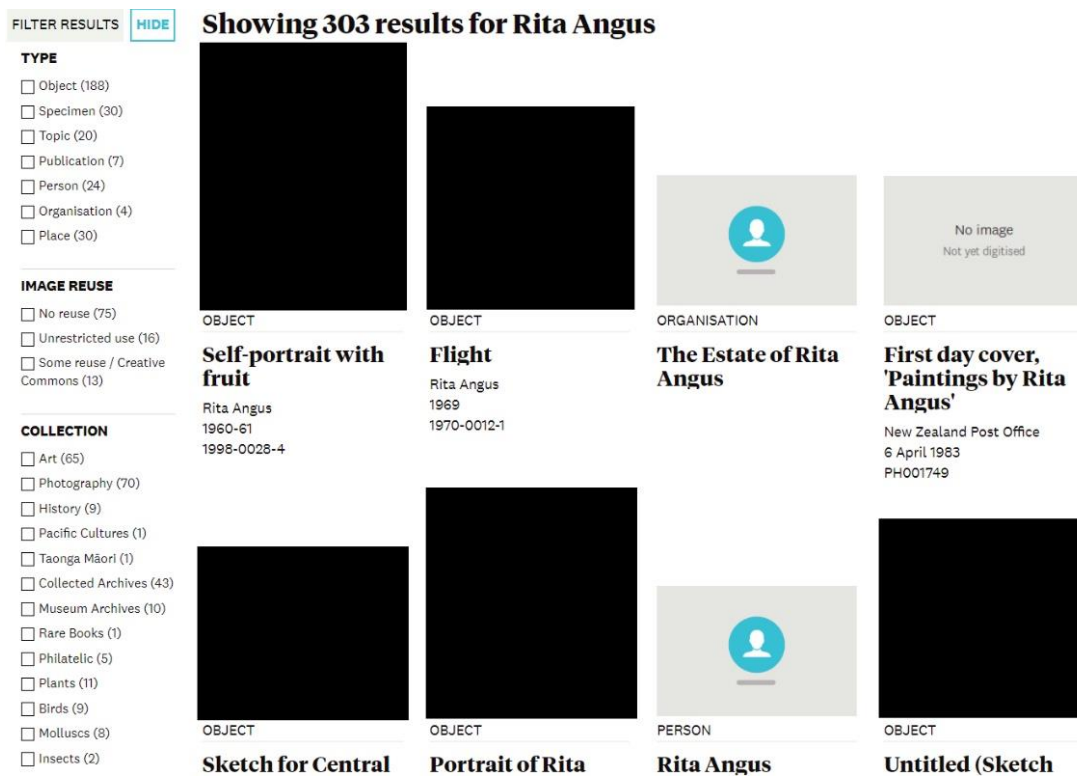


Figure 4: Image of page two of search results on Collections Online, Te Papa for 'Rita Angus,' taken on 07/01/20. (Images of collection objects removed due to copyright).

¹⁵² 'New Collections Online Launches', Museum of New Zealand Te Papa Tongarewa, Wellington, New Zealand, 15 September 2017, accessed 7 January 2020, <https://www.tepapa.govt.nz/about/news/new-collections-online-launches>.

¹⁵³ Fiona Moorhead, 'Te Papa's New Collections Site Launches – Get Hunting', Te Papa's Blog, 14 September 2017, accessed 7 January 2020, <https://blog.tepapa.govt.nz/2017/09/15/te-papas-new-collections-site-launches-get-hunting/>.

Interview participants

My interviewees were all involved in the digitisation process in some form, either directly working on the ACDP or by providing support, putting forward items for digitisation or working directly with the collection management system. They acted as key informants who were able to provide information and insight on digital surrogates, digitisation, and metadata. Acting in different roles they all have different perspectives and areas of interest.

Kirsty Lillico, Collection Manager

Lillico is the Collection Manager for the ACDP. She has experience working with photography collections, particularly after working at the Alexander Turnbull Library on another mass digitisation project. Because of this, she has taken a bit more control and independence in working with the photography collection at Te Papa. She works alongside the two imaging technicians, the collection managers of each area and conservators. Lillico's main role in the digitisation process is the physical care of the objects, retrieving them, delivering them to the imaging team, ensuring they know how to handle them, that they're safe while they're there, and then returning them to storage.

Athol McCredie, Curator Photography

McCredie is the Curator Photography. He started as the Curator Art and Visual Culture in 2001, as there were not specific media disciplines for photography at that point. There are two Curators of Photography, with the other, Lissa Mitchell, dealing more with historical photography. They have a general division of roughly before and after World War I. McCredie broadly describes the role of a curator as selecting things; for exhibition, for the collection, and that everything else they do revolves around this function. A large part of this work is administrative and deals with enquiries related to the collection, or their broader expertise about media and practice. McCredie has been sending items for digitisation across his time at Te Papa and has seen

changes in how this process has functioned. Being focused on photography he holds digitisation in high importance and sees this as a key part of collection management.

Fiona Moorhead, Collections Information System Manager

Moorhead was the Collections Information System Manager during data collection for this research. This role has three main aspects: managing the collections information system at Te Papa, EMu; managing Te Papa's collection related digital assets; and looking after the Collections Online site. Her role in the digitisation process was to provide assistance and advice about technical issues, the import process, the arrangement of metadata, the file types, and troubleshooting of problems. She also received digital collection items and loaded them into EMu, and worked through any issues with them.

Dionne Ward, Imaging Technician

Ward is an Imaging Technician for the ACDP. She describes her role as being involved in digitising negatives, objects, and works on paper. Once she has imaged material, she is responsible for loading it into EMu so that it attaches to the record. She notes that ensuring the object is attached to the correct record is one of the most important parts and is sometimes done quite manually.

The interviews were semi-structured, and based on the interview guide found in Appendix one. Data collected from these interviews and from documentary sources was thematically analysed and the key findings from this data are discussed in the following chapters.

Chapter two: Different worldviews, perspectives, definitions, and domains

When asking about metadata, digital surrogates, and relationships it became clear that the differences in domains that I had found in the literature were also seen at Te Papa. This began with the language I was using which was often used or understood in a narrower sense by my interviewees. Each person I spoke to also carried their own perspective and worldview on the nature of the relationship, the use of metadata and digitisation in general. As I was speaking to people who were already involved with the digitisation process and happy to speak about it, they all felt positive about digitisation and the creation of digital surrogates. In this way, they shared a similar perspective. However, as individuals, they all brought different perspectives and experiences, particularly in how much they considered the relationship between digital surrogates and physical object. In answering my research questions, it became clear that the ideas I had taken from the literature and my experience in libraries and archives, were not spoken about in the same way by my interviewees. This chapter explores this key finding, looking at different definitions and use of language involved in this research, and the priorities and perspectives that impact on the creation of digital surrogates. The drivers of digitisation, access and preservation, are then discussed.

Different worldviews and perspectives

As I mainly spoke to those who were involved or had experience with photography collections a lot of their ideas were specific to this type of media. Everyone within a museum has a different focus, different priorities, and a different perspective on the role and function of the museum, its objects, and what part they play in that. This comes together to mean that different people can view similar things in different ways, depending on how involved in it they are. For my interviews, I spoke to people that sat in different roles to get a small sense of some of this range in perspective. This was important as it allowed me to get a glimpse of how this range of opinions and focuses may contribute to creating metadata, the use of metadata, and how they saw the relationship between digital surrogate and object.

Digital surrogate definition

For this research, I used the term digital surrogate in a broad sense to refer to the outputs of digitisation and digital imaging. This digital asset acts as a digital representation of the object. When speaking about this term with my interviewees, all four expressed ideas which reflect the narrower view of a digital surrogate as being a digital asset that is a true surrogate for an object. None of my interviewees actively used the term, ‘digital surrogate’ themselves. Within the context of my interviews and research, they were all able to understand the broader use of the terms and some would use this language later in the interview to explain certain concepts. Moorhead consistently used the term ‘digital asset’ to encompass digital surrogates as well as born-digital objects, which also reflects how they sit within the CMS.

All my interviewees expressed the idea that the use of the term digital surrogate would better apply to collections like photography, than others. This was because the reason those were collected was the image, rather than the medium. An example of this would be that a glass plate negative is primarily collected for the image it contains and this image is what will be exhibited, rather than the glass plate itself. For something like a painting or three-dimensional object, the medium is more important. For photography, a digital surrogate could replace or stand-in for the original much more effectively. McCredie, the Curator Photography, expressed this very specifically, but this sentiment was reflected by all my interviewees:

I think photography is more important, in many ways, to image than other areas of the collection — I’m sure others disagree — but one of the key factors is that the photograph is an image. Something else in the history collection, like a teacup and saucer, is an object and the image of it is clearly only an image of it, it’s not the thing itself. Whereas an image of a negative is effectively the object. We acquire the negative in order to get an image, whereas we didn’t acquire the cup and saucer to get an image, we got it to display as an object in exhibition.

[...]

The photographs live as image primarily, whereas other things in our collection do not live primarily as images.

[...]

I think that for the cup and saucer it's a digital representation because there are many digital representations for [it]. You can photograph it from above, the side, below, different configurations. These are options for photographs too, but they are much more limited and effectively there's really only one way for it to be imaged and that is a surrogate. (Athol McCredie)

McCredie's role in the acquisition of photography collections is clearly influenced by his perspective on photography and what the content of that format is, as well as the reason for collecting these objects. This broader purpose for the collection is front of mind when considering digitisation. McCredie's point that objects are acquired primarily to be displayed as objects, rather than for their content, like a negative would be, is not repeated by my other interviewees, although both Lillico and Ward also spoke about the greater level of interpretation involved in imaging these objects. McCredie himself notes that others may disagree with his assertion that this means photography collections may be more important to image, but his distinction between the digital imaging of photography collections versus those which are more three-dimensional was spoken about by my three other interviewees who work more broadly across collection formats.

The term digital surrogate was chosen for this research for its broad definition to refer to a digital representation of an object. This sits outside of whether it could be a true surrogate of that object and instead focuses on the use of the visual image, or images, as a proxy for the object. This came out of the libraries and archives literature where most of the material being digitised is documentary in nature, and so the distinction between a surrogate and a proxy is less significant.¹⁵⁴ Museums also have similar types of formats, as well as a greater focus on larger three-dimensional objects, which may account for those I spoke to at Te Papa not using the term 'digital surrogate'. However, needing to capture the experience of some formats is a concern in both domains.¹⁵⁵ Lillico mentioned the concept of artist's books, specifically where the experience of turning the pages is part of the artistic concept of the piece.

¹⁵⁴ Beaudoin, 'Context and Its Role in the Digital Preservation of Cultural Objects', 2; Beaudoin, 'A Framework for Contextual Metadata Used in the Digital Preservation of Cultural Objects', 3; Terras, 'The Rise of Digitization', 3–10.

¹⁵⁵ Anderson, 'Electronic Information and Digitization: Preservation and Security Challenges', 164; Mitcham, 'Emulation for Preservation - Is It for Me?'

She noted that this experience, the feel of it, can be hard to translate when photographing them.

The term ‘digital surrogate’ has a broad definition but clearly invokes a variety of specific ideas when used as a term. All of my interviews involved discussions of this term to get a sense of how the research defined it, but also how they felt it intersected with their work. This was heavily influenced by their own perspectives, particularly the focus on photography collections.

Metadata: catalogue information vs descriptive metadata

The use of the term metadata in this research was also different for each of my interviewees. For McCredie, Ward, and Lillico it sat outside of the information that they contributed to directly, or even use. All three spoke of computer or camera generated metadata, and there was an implication that metadata was the type of information that was captured that provided automatic detail about the creation of a digital image or record, technical metadata, and did not include the information that they found useful, the descriptive metadata.

I think of it as basically digital stuff related to a digital image, as it’s used in photography circles. I never really look at it, but its stuff about what aperture the photograph was taken at and what date and things like that. But I never really deal with that information. [...] It might also mean the cataloguing information about an object, that is what the dimensions of the negative are, a description of it, but I just call that cataloguing information, documentation. (Athol McCredie)

I’m a bit confused by this word metadata, so I always think this word refers to all that stuff that’s embedded in, for instance, a camera, all the settings, and all the other [stuff], but you mean collection information? (Kirsty Lillico)

Both McCredie and Lillico express an awareness of what I am referring to when thinking about metadata, but for them, this is instead ‘collection information’, ‘cataloguing information’, or ‘documentation’. Distinguishing between technical and descriptive metadata in this way seems to be linked to the value placed on it by the user, or perceived user. Creating this distinction, and being clear about what needs resource, which is human-generated, and how it is then used can be important in an environment with tight resources.

This same level of distinction is not reflected in the way libraries and archives write about metadata. Instead, there is a greater focus on metadata of all types coming together to create order and capture information.¹⁵⁶ Moorhead also observed this difference between the domains describing libraries and archives staff as tending to be more ‘database minded’:

I think that libraries and archives often have staff who are much more metadata focused. I think in museums sometimes there is less knowledge around information management and metadata creation than in other areas of the GLAM sector. I think that maybe we are more object focused or maybe we are more focused on creating descriptive content, as opposed to structured content. (Fiona Moorhead)

Compared to libraries and archives, the museum’s focus on being descriptive may also indicate a different reason for collecting, a greater focus on exhibitions, and a more holistic view of the relationships, one which does not have a focus on embedding this information into the record in the same way. This may also mirror differences in arrangement and description.¹⁵⁷ This was reflected in the way my other interview participants saw the structure of the system and the relationship between the digital assets and the collection objects.

Like the term ‘digital surrogates’, ‘metadata’ is a key term in the research. For this reason, ensuring that my interviewees understood what the term covered in my research was important, but also that I understood how they would personally use the term and how they would refer to things which are considered metadata for the research. This led me to an understanding that they would not use the term ‘metadata’ for information and data which would be generally considered descriptive metadata. This was useful for seeing their focus and helped support understanding throughout the interviews.

¹⁵⁶ Duval et al., ‘Metadata Principles and Practicalities’, 10, 15.

¹⁵⁷ Crow et al., ‘A Unique Arrangement’, 335–356; Duranti, ‘The Archival Bond’, 217.

Prioritisation

I was interested in how digitisation work was prioritised to give a sense of why the work was occurring, particularly around access or preservation, in order to answer one of my secondary research questions: Does the purpose, i.e. access or preservation or both, of the digital surrogate impact or effect the way metadata is created, recorded or displayed? While it appears that the purpose does not have an impact on metadata, it instead seems to reflect the differences in formats, and how that affects the way digital surrogates are thought of. Responsibility for prioritisation is confined to select roles within Te Papa, notably curatorial staff. It was noted by Moorhead that priorities have developed and shifted over the project, as one would expect. Due to only speaking to one curator for this research, the information I was able to get about these priorities is primarily related to the photography collection.

McCredie, as a curator, was able to speak at length about prioritisation from his point of view, both for specifically digitisation, but also across the collection and how imaging forms a part of this. For McCredie digitisation is a priority: “I particularly feel like negatives need to be digitised because a negative without an image is really of no use to anybody”. Prior to the ACDP, over a 10-year period, McCredie would put up around 200 objects per year of contemporary work to be imaged. This was sent to a team who also had competing priorities as they had to also work on promotion images for Te Papa and Te Papa Press. Now, as part of the ACDP, curators are asked for priorities to be put forward each year. Even after those initial 10 years of work, there are always still some to do, both new acquisitions and stragglers from previous years. Ultimately the decisions are made within the ACDP, alongside things which pop up more urgently, such as for exhibition. There are also practical difficulties that can arise such as with nitrate film, or when there is no one available to retrieve things from cool store.

Photographs which are currently prioritised for the ACDP fall within a few broad categories for prioritisation. Preservation is a big part of this with nitrate negatives featured heavily, as well as acetate negatives which also deteriorate over time. New acquisitions are also a priority, and with their cataloguing funded for a contractor to

register them, this means they will be accessible online in the near term. Panoramas incorporate a number of digitisation drivers, including preservation concerns, as well as the idea of grouping similar formats together for ease of imaging. McCredie notes, “That’s a case of deciding something is easy to do if you do it all as a group. It also makes access easier and it preserves them”. Other priorities include items which do not have a lot of information about them, and so imaging will make their content accessible, both externally and internally, as well as research priorities for the photography curators.

Prioritisation sits across collection management, however, imaging forms only a part of this. McCredie shared a spreadsheet which goes back to 2002 which uses a matrix to determine priority. The EMu team have also begun using a similar idea. The photography collection matrix looks at significance, public demand, programme relevance, improving access, enhancing preservation, clarifying status, and increasing storage space. These ideas help prioritise collection management activity such as registration, but also digitisation work. For digitisation, ‘how easy will it be to photograph?’ is an important question. Photography tends to be easier than three-dimensional objects because they can use a similar set-up, unlike objects which may be different sizes and shapes.

By looking at prioritisation I was able to gain a better understanding of the key drivers for digitisation, as well as how digitisation sits within other collection workflows and projects. It also supported understanding of the ACDP and the current rapid digitisation occurring at the museum.

Digitisation for access and preservation

Digitisation is often talked about as part of both access and preservation. When asked about these drivers, more broadly across collections, my interviewees spoke about preservation as an early concern which is now only seen with regards to particular items with deteriorating properties, such as nitrate negatives, or other deteriorating formats. Instead, access seems to be the key driver. One of the main focuses of the

ACDP is to image items that do not have existing images. This focus speaks to the idea that having an image for an item increases access to it through Collections Online. This does have challenges, however, particularly where items have been previously imaged but at the standards for that time, such as the Burton Brothers collection, imaged in the 1990s. Those images were still used and accessed by the public but did not reflect today's digitisation standards. However, reimaging this material fell outside of the scope for the project.

I'd rather have an image of any sort than no image, but it's also a question of significance and of use, of interest. I think the Burton Brothers collection is highly significant, it has high public interest and use, and it's our flagship photography collection. We should be looking good with it, not at 1990s standards. (Athol McCredie)

It was eventually agreed by the ACDP team that the low-resolution images in the Burton Brothers Collection would be reimaged, however, these sorts of discussions will continue to be had as digitisation standards and expectations increase.

Preservation is certainly still a concern, Ward notes that while access is a driver, for some objects the preservation element is a key concern with deteriorating formats. However, linking back to the concept of what is and is not a surrogate, for three-dimensional objects the focus is on access. McCredie echoes this and notes that for the photography collection "it's important to get an image, as high a quality image as you can, off that before things get worse", especially for items like large glass plate negatives which are old and very fragile. The focus on preservation appears to have dropped off over time, with a greater focus on access. This possibly reflects both that items with preservation concerns have now been digitised and that the importance and expectations around online collections have increased.

For Lillico, as someone tasked with working with the physical elements of objects, access for her also speaks to physical access and ease of access to retrieve items. McCredie also echoed this point on occasion. This is particularly important when thinking about priorities for digital imaging and whether items are prioritised. Ease is one of the mandates for the ACDP to support efficiency. Public and research access are also drivers, however.

Access is a key driver, and certainly for the ACDP project that has been one of the reasons for receiving the funding. [...] Another one of the decision-making questions about choosing a group of items to digitise is that we can release it on Collection Online. So that's definitely a factor. (Fiona Moorhead)

Considerations for the physical access to collections for staff contrast with the digital access for the public that digitisation provides was only touched on by Lillico and McCredie, who both have roles which focus on the collection objects and access them in both ways. This indicates a greater awareness from those who are retrieving objects from storage around the access implications that exist physically. Certainly, the ACDP started with a focus on digital access without accounting for the physical access that would be required. A collection manager for the project was not funded from the outset but was later included as a required resource.¹⁵⁸

Moorhead's awareness of the duality of access and preservation is also interesting in thinking about the specific formats that this relates to. This is reflected in which images have a preservation master created for them (those which could act as true surrogates, such as negatives), and those which only receive access masters. Preservation masters are high-resolution, unedited files which are stored in a separate repository, accessible by only a single staff member for security. An access master is also created which is a lower resolution, but still high-quality, and may include relevant edits, such as reversing the tones for images of negatives. Derivatives, or files of various sizes and types, for general use are created from the access master.

It's pretty clear that the organisation benefits from the digitisation both in terms of access but also preservation because it's going to be harder to use those [nitrate negatives or glass plate negatives] physical items in the future. (Fiona Moorhead)

Even when thinking about physical access and preservation this duality exists. Lillico spoke about locating and sighting items when retrieving them for digitisation, which provides opportunities for brief condition checks and reconsidering storage.

¹⁵⁸ Athol McCredie (Curator Photography, Museum of New Zealand Te Papa Tongarewa), interview with Laura Jamieson, October 2019.

The concept of digital access was often referenced by all interviewees and was definitely the driving focus in the conversations around digitisation and its value. Certainly, public access was key, as was having a large amount of the collection available. McCredie's focus on photography also reflected this idea. He noted that photographs tend to come in volumes, especially in museum collections. McCredie provided the example that when dealing with 10,000 negatives, as opposed to 10 teacups, not every image will be that important, but you need to see them all to have a useable collection. He notes that for photography, institutions tend to create image libraries, so when discussing access, it needs to be thought of to a different standard.

Access and preservation are the key drivers for digitisation across institutions and this is reflected in the literature and this research. Hearing about these drivers from people who work across digitisation, helped to gain a more nuanced understanding of what access and preservation look like at Te Papa. Comparing this against the literature led to a recognition of a difference in the way access, in particular, is thought of and approached differently in this museum context than is explored in the libraries and archives literature.

Different domains

This distinction between the different domains; museums, libraries, and archives, and ways of thinking was most evident when asking my interviewees about access and preservation in relation to digitisation and the creation of digital surrogates. For libraries and archives, access and preservation in digitisation go hand in hand. The objects they have in their collections are physically accessible and digitisation allows for remote access and greater discoverability.¹⁵⁹ This also supports preservation as

¹⁵⁹ Ciurea and Filip, 'New Researches on the Role of Virtual Exhibitions in Digitization, Preservation and Valorization of Cultural Heritage', 31; Conway, 'Digital Transformations and the Archival Nature of Surrogates', 55; Henning, 'New Media', 306, 309; Jardine, 'Reflections on the Preservation of Recent Scientific Heritage in Dispersed University Collections', 741; Reynold, 'The Digital Initiative in Archives', 19, 27–29; Robinson and Tanner, 'Higher Education Digitisation Service: access in the future preserving the past -- the UK perspective', 66.

digitising an item can reduce the physical handling of that object in the long term.¹⁶⁰ Some institutions stop physical access once the item has been digitised, but some also do not, and for formats like maps the ability to see something physically remains important. Digitisation still provides a level of access to that item and can also provide new information through things like zoom and being able to digitally overlay.

With museums, this type of physical access by the public does not happen in the same way. Objects in the collection may serve a research purpose, but this tends to be limited and specific. Objects are not routinely retrieved to be accessed by anyone who walks in. Physical access to objects tends to happen in the exhibition space, where this physicality is mediated by protective measures. Objects in museums are tangible items and, as Parry discusses, their materiality is important to their authenticity for the public.¹⁶¹ Digitisation for preservation in museums is then only referring to items where the digital image could be a surrogate, such as for photographic images, where the digital asset could replace the deteriorating physical object. Otherwise, the idea of preservation is much more about general conservation principles of seeing an object over time and being able to compare its condition visually. This also is not a primary driver for digitisation.

Access is this primary driver, but it is a different way of talking about access. In talking to my interviewees, it strikes me that imaging is discussed as the be-all and end-all of access. That providing a visual representation of the collection object online is making it accessible. Access is not to the physical object, but rather to the information about the object, including visual. In this way, the images of objects could be considered to all be digital surrogates, not just those which could function as true surrogates of content. Items within a library or archive become accessible through discoverability, their content may be accessed digitally, but equally, without being digitised they are often still available to be accessed. For museums, this access

¹⁶⁰ Ciurea and Filip, 'New Researches on the Role of Virtual Exhibitions', 30; Conway, 'Digitizing Preservation', 45; Gould et al., *IFLA/UNESCO Survey on Digitisation and Preservation*, 26–27; Mani, 'Digitisation: Preservation and Challenges', 70.

¹⁶¹ Parry, *Recoding the Museum: Digital Heritage and the Technologies of Change*, 61, 64, 68, 76.

is much more mediated, and digitisation becomes the primary point of access to collections that are not on display.

The common misconception of museum collections hidden in basements and statistics about the percentage of museum collections that ever go on display speaks to this lack of access to a museum collection that does not exist in the same way in libraries and archives.¹⁶² While my interviewees also speak of discoverability, for them Collections Online is the portal for access. This is how things in the collection are able to be seen. The desire for these things to be seen and accessed is clear when speaking to my interviewees, and that this approach has Collections Online as central to providing access. This is also considered adequate for public access, perhaps because museums see this working alongside their exhibition focus to provide more complete access. It may also speak to differences in collecting priorities and drivers between museums, libraries, and archives, as well as differences in institutional mandate.

Summary

This chapter has explored the different worldviews, perspectives, definitions, and domains present in my research. While I set out to speak to people in different roles in order to gain a wider range of perspectives, it is also clear that these perspectives are just a small part of the much wider range that may be present at the institution. Ways of thinking about digital surrogates and metadata depended on the type of interaction a person had with it and how they were used by them. Equally the type of language used in the literature would appear to have come out of libraries and archives and so does not always hold the same meaning or value within this museum environment. Being aware of this, and accounting for it in my findings is important, but it also indicates a different way of thinking about digital surrogates or metadata which reflects the differences between domains.

¹⁶² Christopher Groskopf, 'Museums Are Keeping a Ton of the World's Most Famous Art Locked Away in Storage', Quartz, 21 January 2016, accessed 26 February 2020.

These domain differences also extend to the drivers behind digitisation in the first place and the reason digital surrogates are created. What came through in analysing the interviews was that the concepts of access and preservation hold different weight within the museum domain, which had different structures around access and different focuses in preservation. Ultimately, in attempting to answer the research questions, what was clear is that the approach to digital surrogates, the relationship and metadata at Te Papa was different to what I may have expected coming from a libraries and archives background. However, this difference in approach still shared similar fundamentals and gave me an insight into the use of the system architecture to represent the relationship, rather than a focus on using metadata.

Chapter three: System architecture representing relationships

When answering the research question, ‘in what ways is the relationship between physical collection items and their digital surrogates represented using metadata at Te Papa?’, it became clear that while there are instances where the relationship is represented by metadata these are largely incidental representations (See Chapter four). This is because the system architecture is driving the relationships. While I asked about the use of metadata to represent the relationship between the digital surrogate and the physical object, it was revealed that the use of metadata is influenced by the way the system is set-up. The CMS is structured in such a way as to embed these relationships within it. The collection record sits at the top of the hierarchy, functioning as a parent record to which multimedia and media asset records are linked. For this reason, the relationship is obvious and intuitive to users, both internal and external. Metadata then becomes less important for representing this relationship. This chapter explores this key finding, by looking at the system architecture and how it is structured beginning with records and how they are related to each other, the differences between types of metadata, and finally how the system, the metadata, and the relationships are linked to the Collections Online interface.

System structure and relationships

The collection management system, EMu, is structured with three types of records, two of these descriptive. The main record, which is what can be searched, is the collection record, which provides descriptive metadata for collection objects. Media asset records provide descriptive metadata for both analogue and digital media assets which relate back to collection records. They contain information which describes the media, not the object portrayed in the media.¹⁶³ These may then have multimedia records, which just have the multimedia file attached to them, and are linked to the collection record. Each multimedia record must be attached to a media asset

¹⁶³ Kingston, ‘KE EMu Media Assets Manual, General Staff Version’, 12.

record.¹⁶⁴ This structure is seen in Figure 5. For analogue media assets which are digitised, the analogue asset will have a media asset record about its properties ('Analogue Media Asset Record' in Figure 5) and then a media asset record will be created as a child record to this for the digital asset ('Digital Media Asset Record (i.e. scan of Analogue Media)' in Figure 5), which will then have the multimedia attached to it. There is a clear distinction that media assets are not collection objects, but also that they are being managed over the medium or long term.¹⁶⁵

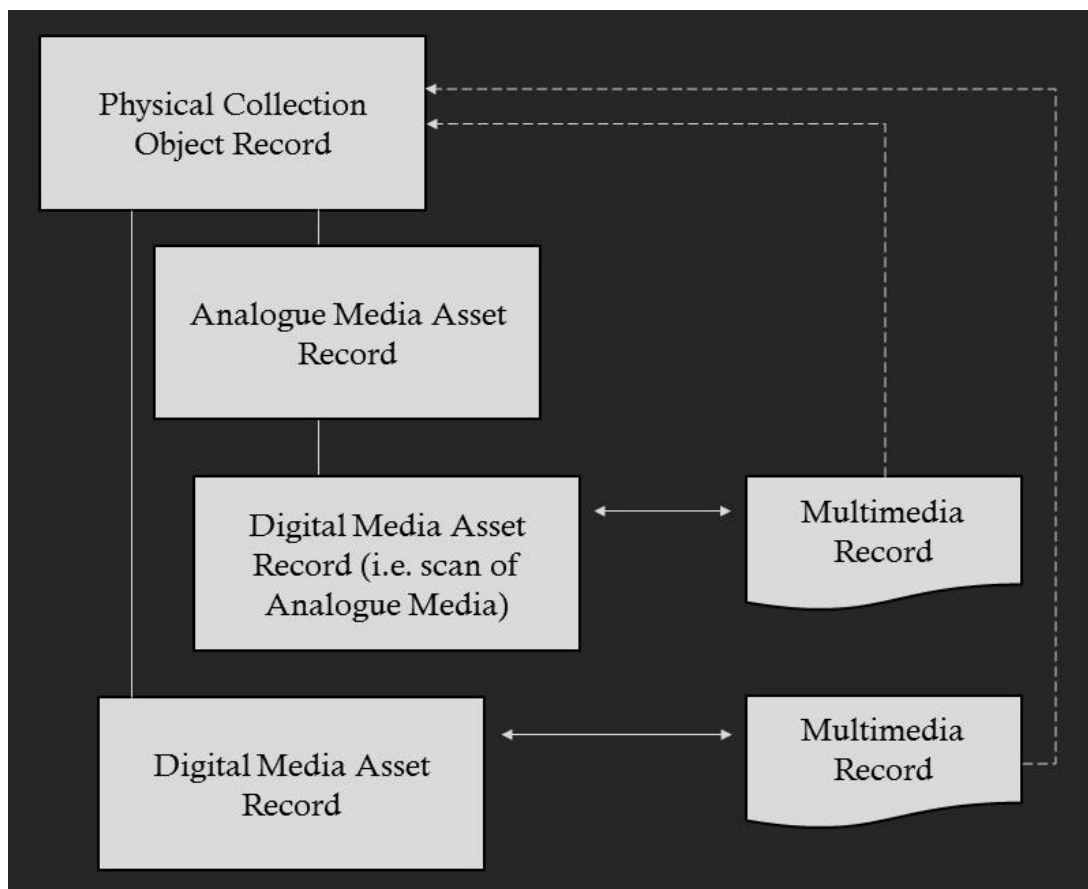


Figure 5: EMu record structure. Simplified diagram adapted from Museum of New Zealand Te Papa Tongarewa, KE EMu Media Asset Manual, General Staff Version, 4.

Moorhead expressed that this type of system has been complicated for some staff to understand, and my other interviewees did not find the detail of the way the system

¹⁶⁴ Kingston, 'KE EMu Media Assets Manual, General Staff Version', 3. However, for born-digital objects the multimedia will instead be attached directly to the collection record.

¹⁶⁵ Kingston, 'KE EMu Media Assets Manual, General Staff Version', 3.

was structured to be relevant to the work they were doing. They were instead more focused on the specific part of the system their job required them to interact with. Most users interact with this system from the collection record. From here, multimedia can be seen in a section of the record (see Figure 6), and information about these media assets can be accessed from this record if needed.

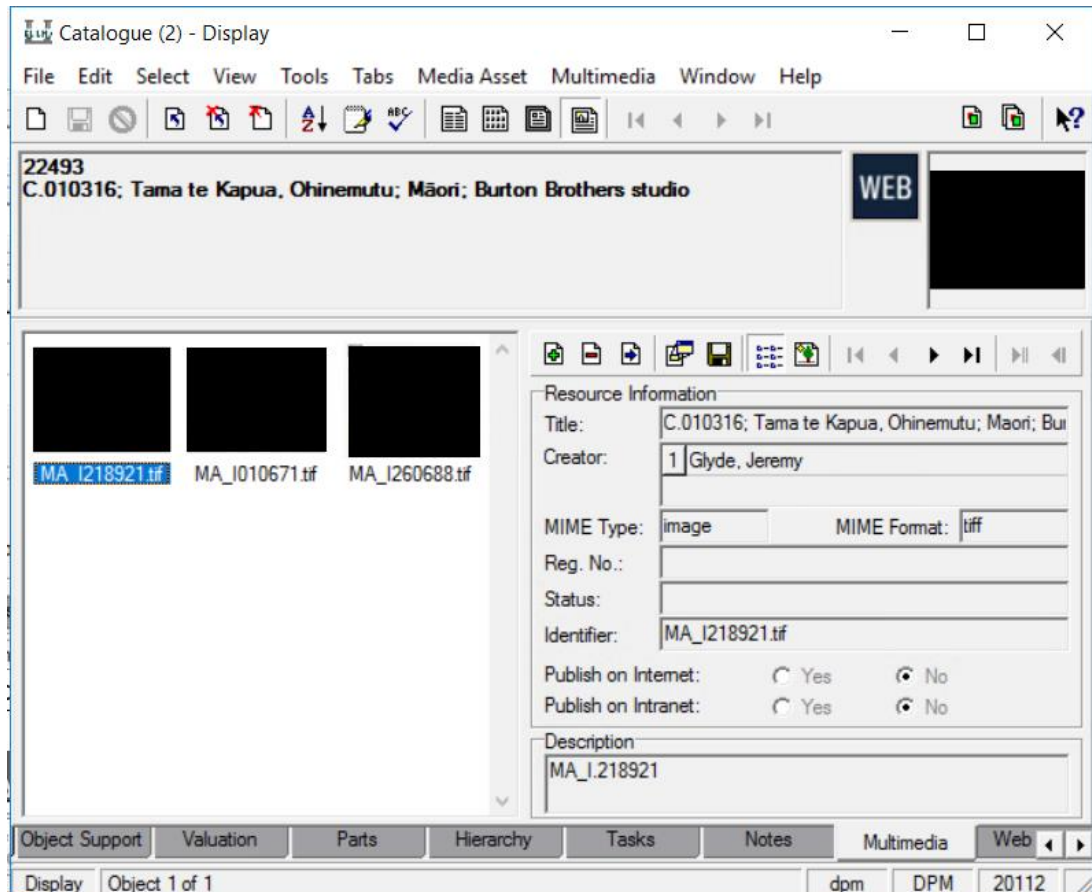


Figure 6: Collection object catalogue record (IRN 22493) in EMu, showing multiple multimedia as thumbnails. (Images of collection objects removed due to copyright). Taken from 'EMu Media assets at Te Papa: an overview of the current state, Museum of New Zealand Te Papa Tongarewa, March 2019, 8.

Both Lillico and McCredie did not generally go into the media asset record to find the information captured there, such as the photographer, the file size or type, or information about the equipment used for the digital image. For them, the information about the collection object itself is much more important. Ward has more familiarity with the information in the media asset record as she creates it in her role

when uploading the images to EMu. For Lillico and McCredie, the digital media itself, the image, functions as a descriptive element of the collection record. Despite this focus on the collection record, this way of setting up the system is “very pure” in terms of digital preservation and provides a good foundation for caring for these media assets as part of the wider collection.¹⁶⁶

The focus on the collection record does mean, however, that the relationship between the collection object and its media assets, both analogue and digital, is mainly represented by the CMS itself. Often the relationship is seen as obvious and straightforward because in accessing information a user can immediately link the two together. Without a clear understanding of the way the system is structured, nor why, a user is not required to think much about the relationship, how it exists, nor whether it could be represented in other ways.

This affects how my interviewees saw the digital asset too. In a similar way, as it is structured in the CMS, there is a parent/child relationship between the collection object and the digital asset. The digital asset’s value is directly related to the collection object, providing access, some preservation of content, and information about the object in time. The collection object has value without the digital asset, but the digital asset does not have this same value as a standalone item. Despite this, they do still have some value to the organisation, and this is reflected in the fact that media asset records, including multimedia elements for the digital media assets, are retained over time. This ensures that they serve as a record for several different things. This includes information important to the collection item, such as its condition over time and the processes it has been through regarding imaging, as well as keeping outputs from the organisation and recording imaging processes over time including tracking the percentage of collections imaged over time.¹⁶⁷ Being used in this way also shows the importance of the relationship back to the collection object, and how the digital surrogates have value based on this relationship, reflecting the

¹⁶⁶ Fiona Moorhead (Collections Information System Manager, Museum of New Zealand Te Papa Tongarewa), interview with Laura Jamieson, October 2019.

¹⁶⁷ Fiona Moorhead, via email correspondence with Kirsty Lillico, November 2019.

ideas behind archival arrangement and description.¹⁶⁸ Again, however, this relationship happens within the CMS interface, with access to the digital asset coming from the collection record.

This again highlights how the CMS, and the way it has been set up and is used, as the primary way for these assets to be linked back to the collection object, and to each other. Information stored in the individual media asset records do not refer to one another, the existence of each of these media assets is only clear when looking back at the collection record. This is illustrated in Figure 7 which shows this record hierarchy with the collection object record the only level containing links to both media assets. This shows how it sits as the parent record and through this the variety of multimedia can be seen, compared, and their media asset records accessed. That this comparison happens from this parent record, means that the importance of the collection object is highlighted, and the record is improved by the inclusion of the multimedia.

¹⁶⁸ Crow et al., 'A Unique Arrangement', 335–356.

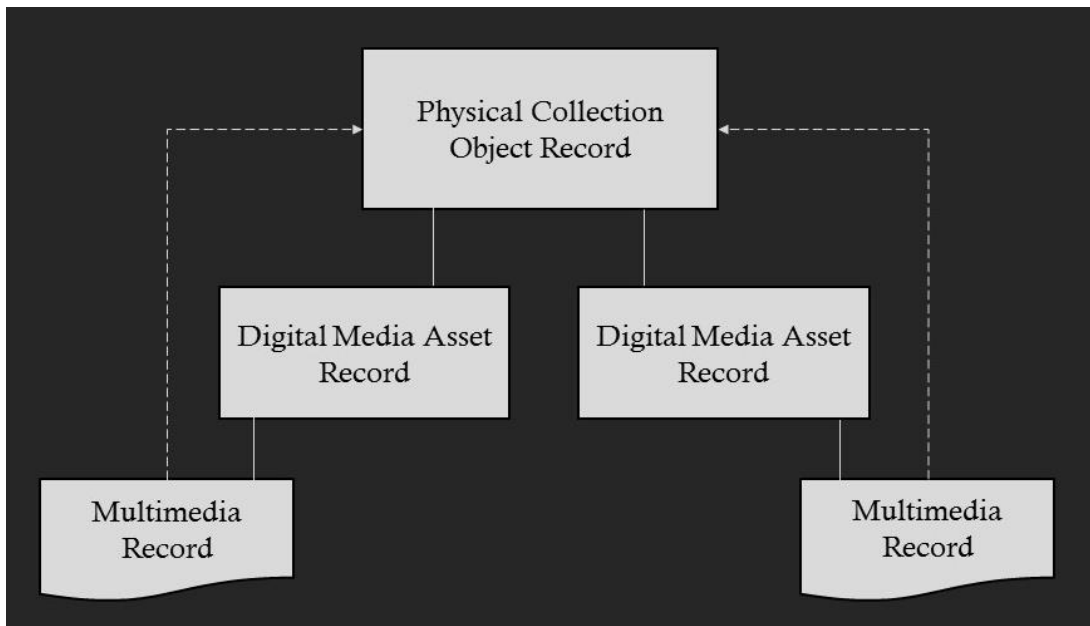


Figure 7: EMu record structure showing the hierarchy of records. Demonstrates that related media asset records can only be seen in the parent record; the collection object record. Adapted from Museum of New Zealand Te Papa Tongarewa, KE EMu Media Asset Manual, General Staff Version.

Even with this focus on relationships, the relationship between the digital surrogate and physical object is collapsed in this view. The collection record information is the only information visible, with the multimedia available but without its information from the media asset record, as seen in Figure 6. Again, this metadata is deemed less valuable to the user. The digital asset, not itself part of the collection, but rather a supporting piece of visual information for the collection record.

Technical and descriptive metadata

The use of metadata in the CMS can be thought of in two main ways, technical metadata and descriptive metadata. Much of the information in the records, both for collection objects and media assets, relates to the physical properties of the item (material/format, dimensions/size). There are a number of mandatory fields when creating a humanities record; registration number, object classification, title, maker, materials, credit line, measurements and storage location, and other fields are

mandatory if the information is known, including production place and date.¹⁶⁹ Other information that can be captured in collection records includes information about the content, the techniques used, what is depicted, and possibly how this item relates to people, places, or events. Much of this descriptive information can be found in the subjects and associations fields, but free text descriptions are also used. The subjects and associations fields are considered ‘mandatory if applicable’.¹⁷⁰ Information contained in subjects and associations fields of a collection record is not replicated in the media asset records, again highlighting the primary nature of the collection record. The media asset record only describes the media asset and not the collection object and it is explicitly stated in Te Papa’s humanities cataloguing guidelines that “the information describing the object should be in the object record”, rather than captured or duplicated in the media asset record.¹⁷¹

While the creation of collection records is a manual process, descriptive metadata is added to the digital media asset records in a bulk ingest process. Ward does this as part of her role as an imaging technician:

A lot of it is automatically generated as I fill out the fields in a bulk loading interface in EMu. The images files are automatically married up with their catalogue record. This feature was developed for this project (ACDP) because we needed to be able to upload so many media asset records at once. (Dionne Ward)

Bulk loading involves filling in which area of the museum the collection item is in - which collection, the type of collection object, the standard - i.e. access master, the name of the technician and role, and that it is made by the imaging team. Then it is tagged with the project, ACDP, the date it was produced, and where it is going (such as the preservation repository or for access). This process ensures a minimum standard for the media asset record.

Much of what I have described so far is descriptive metadata, which was what three of my interviewees referred to as ‘collection information’, ‘cataloguing information’,

¹⁶⁹ Adrian Kingston and Carol Stevenson, ‘Core Cataloguing Guidelines – Humanities Collections’ (Museum of New Zealand Te Papa Tongarewa, Version 1.2, July 2015).

¹⁷⁰ Kingston and Stevenson, ‘Core Cataloguing Guidelines – Humanities Collections’, 25–28, 33–34.

¹⁷¹ Kingston, ‘KE EMu Media Assets Manual, General Staff Version’, 22.

or ‘documentation’. Technical metadata, or automatically generated metadata, is captured in other ways. An example of this might be EXIF data, as Ward explains: “The EXIF data is stored with the file in KE EMu. The EXIF data contains mostly output data, has been processed by the software and some information about input, (camera specifications)”. This type of information is rarely accessed, usually for very specific purposes by a select few internal staff. It is downloadable from the media asset record, but not searchable. Of note is that the camera operator is not captured in this information because they swap and share cameras.

As discussed earlier, this distinction between types of metadata, whilst still recognising them both as such, is common in libraries and archives literature. Using different language to describe some types of descriptive metadata may indicate a different focus or an environment where the use of the term metadata indicates a non-human-generated, and therefore low resource, type of information. A wide range of metadata is created to sit alongside collection objects at Te Papa and this information, whether referred to as metadata or not, records the context of the objects and some of the relationships they have. In discussing both descriptive and technical metadata, the way in which this reflects the hierarchy of the system architecture is clear, with the more labour intensive human-generated metadata sitting with the parent collection record, and descriptive metadata, from the bulk ingest process, or technical metadata sitting in the media asset record.

Collections Online and fuzzy search

Accessing the information and metadata captured in these records is an important part of its creation and was addressed by my interviewees when talking about Collections Online. This public interface was used by McCredie internally for two reasons. The first is for loading capability, especially for records with multiple photos. The second was for its search. Collection Online uses Elastic Search which is a fuzzy search and allows for keyword searching across the collection.¹⁷² McCredie

¹⁷² Fiona Moorhead (Collections Information System Manager, Museum of New Zealand Te Papa Tongarewa), interview with Laura Jamieson, October 2019.

notes that this is particularly useful when you do not know which field a keyword may have been put into.

You don't necessarily know how people have catalogued things in the past because standards have changed over [time].

[...]

If you search for 'house' somebody might have put it in the title, or it might have been in the description. It could be in associations tab, could be several different places and you don't know, whereas Collections Online will just search the whole lot. (Athol McCredie)

Changes in metadata creation over time affect the ability to search within EMu and may mean that attempts to see represented relationships in the metadata are complicated by different standards. The difficulty in performing this type of search in the CMS means Collections Online allows for greater flexibility in finding items that link to each other.

Moorhead notes that work has been done to adjust the relevancy sort for search results to make sure that people find what they are looking for. This manual work to improve searching is also an interesting factor, as it still uses the same collection information, but with other input which may bring out interesting relationships within search results. However, this search functionality could always improve. One of McCredie's visions for the future involves an improved Collections Online search. That this is something that is top of mind for McCredie in a discussion around relationships indicates how important this searching is to maintaining relationships between objects in a collection. What this also shows is the way in which the structure that this information sits in, whether the CMS or the backend of the Collection Online website, has an impact on how those relationships are viewed by the user. For McCredie, searching for specific things requires the use of a Boolean search to ensure the exact thing being looked for comes back. For a public user, a broader range of possible things may be the goal. This wide spectrum of need when keyword searching a large collection highlights the ability for relationships with metadata to be important within the collection record.

Collections Online only presents the collection records, with limited information about the digital media asset, found once inside this record. In this way, it mirrors the hierarchy and structure of EMu and again makes the relationship between the digital surrogate and collection object obvious and clear. Ultimately the collection record is again seen as the primary record, with its metadata that which is searched. For the public user, the relationship between digital surrogate and object is collapsed to the point where they are seen to be the same thing. As stated before, the digitisation of the museum collection object is what provides the public user access to that item. For a user of Collections Online, the digital image or images of that object are integral to the collection record and act as a visual representation of that item.

Summary

While I set out to ask about the use of metadata to represent the relationship it became clear very quickly that this was not the primary way it was represented. Instead, the questions I was asking about the relationship led me to understand that the system architecture and structure held this relationship. Through the hierarchy of the system, the collection record is the primary point of interaction for both an internal and external audience. For this reason, those I spoke to tended to think of the relationship as obvious and clear, despite little metadata which represented it. Instead, the way that multimedia is embedded into the collection record, as well as the way media asset records are accessed internally provided a strong and obvious way for digital surrogates to be related back to the source material. The distinction between technical and descriptive metadata also mirrors this hierarchy, with more descriptive and labour-intensive metadata sitting in the collection record and simpler descriptive metadata and technical metadata sitting in the media asset record. Collections Online also follows this hierarchy, with digital surrogates seemingly embedded in the collection record information and forming a key part of it for users.

Chapter four: Types of metadata representing the relationship

While much of the relationship between a digital surrogate and physical object is represented and held within the system architecture, my focus in questioning was on metadata. For this reason, I was able to ascertain some of the ways metadata represented this relationship, including fields which did this as an incidental function. These are metadata fields which are primarily used for other reasons, but which may also imply the relationship. While this implication is there, my interviewees only considered these fields relevant to the relationship during my interviews. For all four, the relationship was best described through the system architecture and obvious, but when pressed for metadata which represented the connection, they were able to provide examples which may imply the relationship exists. However, because these types of metadata have a different primary function, and this relationship is not front of mind for those I spoke to, it is possible that these fields may be vulnerable, to change or removal. This would mean that any benefit to retaining the relationship between digital surrogate and object through the metadata could be lost. In this chapter on this key finding, I will explore the use of metadata for broader object relationships, including the human element of this, as well as the focus on discoverability and policies which support this. I will then look at types of metadata which do represent the relationship between a digital surrogate and the source material, such as visual metadata, location history and its incidental role in this, and the documentation of interpretation and choices in the digitisation process.

Metadata and object relationships

Subjects and associations are the primary way that relationships between collection objects are represented, through their link to people, places and events. When discussing relationships in my interviews all four people I spoke to brought up these fields. These fields seem to be the way meaning, context, and relationships are maintained in the record.

The collection records contain subjects and associations fields which use standard thesauri to represent the relationship different objects have to one another, as well as to particular figures, ideas, or places. This is the primary place that any sort of relationship is seen. All interview participants brought these fields up as their primary point for thinking about relationships. When a term is used in these fields it then has a mandatory field which requires a description of the relationship, such as maker, depicts, or similar. This centralises the idea of relationships into the metadata of the record. Lillico notes that her own registration practices have changed over time, particularly in relation to what allows discoverability and access to an item:

For instance, I used to write long descriptions for things. But I now know that that doesn't help access or discoverability, whereas now I would use a whole lot of keywords in the subjects. Then they are going to make it discoverable. So, my emphasis has kind of changed in that sense. (Kirsty Lillico)

This indicates the importance of these fields and these relationships to both staff and an external user of these collections. This also indicates a change in behaviour to be more 'database minded' and may reflect a changing culture in thinking about these items, something which Moorhead expressed a desire for.

Collections Online also uses the subjects and associations fields to link together items and provide context to figures. Moorhead describes the importance of these links, and how these are one of the fields hyperlinked on the site. Equally looking at the page for an individual would help them distinguish, using the information contained in subjects and associations, collection objects that person created, and objects they are depicted in. The importance of this is clear to some within the organisation, but the online interface itself, while having this functionality, does not make it a priority for the user. A small drop-down box, seen closed in Figure 8, indicates that 'all' collection objects are shown and requires a user to understand what type of information they could get by clicking on it (see Figure 8a (insert)). Even before this navigating to the 'person' page for a subject they are looking for requires a focus on that subject that many users may not have. A straight search for 'Rita Angus' does not contain her person record on the first page of search results (it appears near the top of page two) (See Figure 4).¹⁷³ Adding Boolean quotation marks

¹⁷³ Search for 'Rita Angus' at <https://collections.tepapa.govt.nz/>, accessed 10 January 2020.

brings it to the top row of page one. The ability to see and use some of the information captured in the subjects and associations fields exists on the online interface, indicates some awareness from the institution of the importance of differing relationships to an object, and how this may be useful for their external audience.

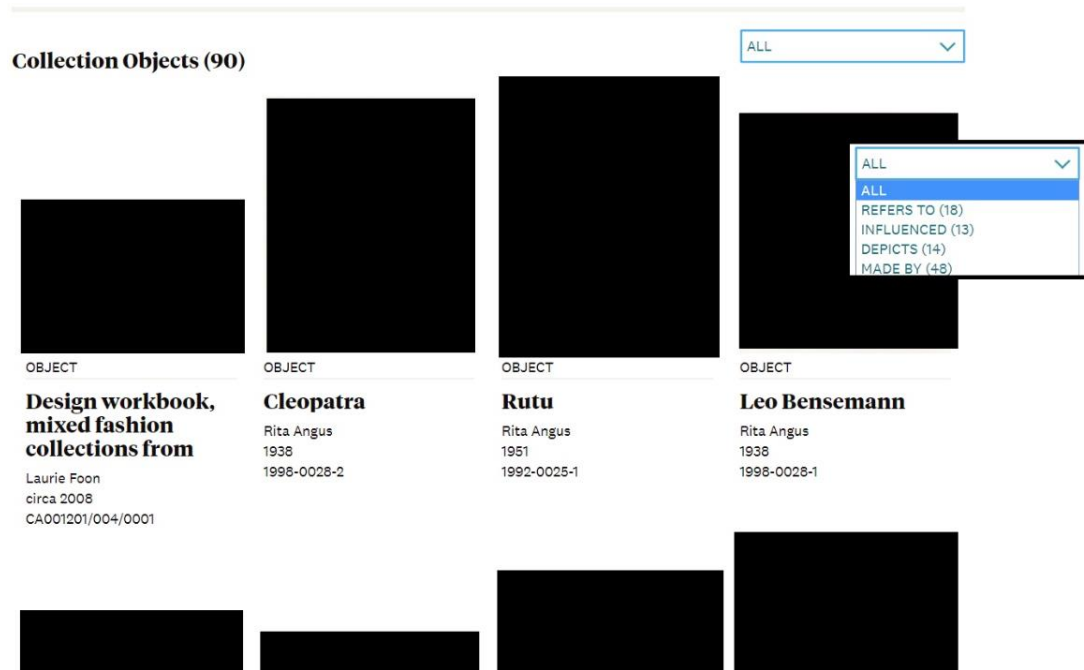


Figure 8: Image of 'Collection Objects' on Rita Angus Person page, Collection Online, Te Papa, <https://collections.tepapa.govt.nz/agent/74>, taken on 3 February 2020. (Images of collection objects removed due to copyright). Figure 8a (insert): Expanded menu showing filter options.

The human element

Moorhead also describes how external web developers often place a higher importance on these links when they discover they are human-generated. These links are then seen as important and also crucial to the context of a collection object within the larger collection or within the landscape or society that the object was collected from.

The idea of something having more value because of the human element is an interesting point to make when discussing links and relationships between items. The descriptive metadata is often human-generated and therefore has a higher value to the user. This relates back to my interviewees defining metadata as that which was automatically generated and not something they interacted with. Instead, they placed a higher value on collection information, which is descriptive and human-generated, possibly because of the effort and manual process involved.

McCredie spoke at length about the manual process that reconnecting items takes, using the example of the Leslie Adkins project, connecting negatives, prints in albums, and his personal diary entries. The Leslie Adkin project focuses on George Leslie Adkin, a “Horowhenua farmer, photographer, diarist, and self-taught geologist, archaeologist, and ethnologist”.¹⁷⁴ It initially involved volunteers transcribing his digitised war-time diaries.¹⁷⁵ The most recent iteration had an intern “matching negatives held in Te Papa’s collection with prints in photo albums that belong to the Alexander Turnbull Library”, as well as working across the negatives, photo albums, and diaries in both institutions.¹⁷⁶

While the relationship between these items is known within the institutions, the work to link these in the system is a manual process, either by linking records or by referencing information across records. EMu does not automatically backlink records in this process either, however, this highlights again the importance of the CMS structure to the relationships between objects. With this example, it is also interesting to note that this linking between collection objects is about context and providing this information to the image, or the image to the information. As discussed in the literature, context is important for how an object is understood and received, and this

¹⁷⁴ Fiona Moorhead, “‘Good-bye, moon’: completing the Leslie Adkin diary project”, Te Papa’s Blog, 21 December 2018, accessed 3 February 2020, <https://blog.tepapa.govt.nz/2018/12/21/completing-adkin-diary-project/>.

¹⁷⁵ Moorhead, “‘Good-bye, moon’: completing the Leslie Adkin diary project”.

¹⁷⁶ Danielle Campbell, ‘Making matches: Reconnecting Leslie Adkin photographs’, Te Papa’s Blog, 31 January 2020, accessed 3 February 2020, <https://blog.tepapa.govt.nz/2020/01/31/making-matches-reconnecting-leslie-adkin-photographs/>.

This project was in the planning stages when my interviews took place.

can be seen in the Leslie Adkin project which prioritises these relationships in order to provide each object with better context.¹⁷⁷

The Leslie Adkin project was interesting for this research in particular because of the use of digital surrogates. McCredie noted, “It’s about pulling those things together and doing that electronically”. The negatives, which have very minimal information with them, have been digitised. These images are then much more accessible for the internal researcher to be able to manually match to the album prints which contain additional information. The album pages have been digitised, but not the individual prints contained within. This complicates the project as these images are attached as multimedia to the album record. However, since this digitisation work was done, registration of the individual prints has started to occur. These records do not have images attached. Whether the album pages are attached to the individual records remains to be seen. Regardless, there will then be a matching process to compare the digitised negative image to the digitised pages of the album to see the corresponding print. Once this match is made this link can be recorded in both records and new information about the image found in the album, particularly the date, can be used to find the corresponding diary page, also digitised and many transcribed, in order to find more context, and add information to all three records. This has many parallels to Jones’ work on the field diaries, looking at items all in the same institution and reconnecting them.¹⁷⁸ It remains to be seen if adding this information to each collection record in the Leslie Adkin project achieves Jones’ vision of a connected collection, particularly in the online interface.

Discoverability

Many of my interviewees spoke about the idea of discoverability, often when referring to Collections Online. The ability to be found for collections was important for my interviewees so that users can find what they are looking for. However, an interesting point was made by more than one participant about the role of digitisation

¹⁷⁷ Unsworth, ‘The Value of Digitization for Libraries and Humanities Scholarship’.

¹⁷⁸ Jones, ‘Documenting Artefacts and Archives in the Relational Museum’, Chapter 3.

in this process. There was a sense of frustration that resource was going into digitisation but not into cataloguing work, which is what allows things to be found.

My vision is that a substantial amount of the photography collection is imaged and available to the public.

[...]

Without good keywords, it's not really more findable than it was before [being imaged]. (Athol McCredie)

This highlights the importance of metadata creation and recording to the digitisation process, particularly for discovery. The metadata is what is searched, both on Collections Online and EMu, and the information it contains provides the context for that object which allows it to be found. The desire for keywords facilitates this searching while also providing a clear standard for metadata creation.

While Lillico has changed her approach to registering items to one which focuses more on subjects and associations than a lengthy description field, she sees value in both.

The descriptions are still useful internally, especially for imaging technicians if they are confused. Sometimes there will be a run of negatives and they're all pretty similar, sometimes they can look at the description and tell the difference if something goes wrong.

[...]

Thinking about access, definitely association, people, places, subjects, they're all really useful. (Kirsty Lillico)

Metadata that you might expect to provide differentiation between objects, like the title, may not provide this either. Lillico notes that for collections like photography there is a hierarchy for creating titles, starting with what is inscribed on the photo or an existing title provided by the photographer. These titles may apply to a group of negatives, however, and so may not be descriptive for the individual images. For this reason, having both the free text description and the subjects and associations can help increase discoverability.

Policies, standards, guidelines, training

I identified documented policies and standards in my research questions as an area of interest, as well as whether they reflected practice. What I found is that while policies and standards for creating metadata exist, including catalogue information, their primary method of transmission to staff is through training. Of course, behaviour can be different across the wide range of people who work on records. This research looked at the cataloguing guidelines, both for media asset records and for humanities collection records. Due to the lack of metadata representing the relationship between digital surrogate and physical object, there was little discrepancy between the documented process for the creation of metadata which does this and with what was occurring in practice. To support consistency across records more generally, there are some mandatory fields to ensure minimum standards, which are signified using green. Equally the subjects and associations fields have mandatory relationship signifiers, as well as standard thesauri.

Specific subject areas may also have their own thesauri used for other fields. Art, for instance, has had its own standards for many years. High staff turnover, however, means that current staff may not all be up to speed with this. The photography team also have developed their own standards, including a list of terms for media which is important for how something is described. This is used in the media summary line which is a free text field.¹⁷⁹

Examples of metadata representing the relationship

When the focus is on intentionally representing the relationships objects have to one another, the link between the object and its digital asset is less prioritised. As discussed previously, much of this is because this information sits within the CMS and so it appears obvious to the user that these two items are linked. However, due to the hierarchal nature of the relationship, the main locus of these links is the collection records. This means that from that end the relationship is obvious to the user, from

¹⁷⁹ Athol McCredie (Curator Photography, Museum of New Zealand Te Papa Tongarewa), interview with Laura Jamieson, October 2019.

the media asset record this link is less clear. Media assets are named by pulling in information from the collection record, the title, creator, registration number for instance. This then provides some way to link back to the original. The use of registration numbers for maintaining relationships has been part of museums for some time and can be used to reconnect objects in the event of a disconnection that is noticed.

Metadata is, of course, a broad term and extends outside of information captured in the record. Information about a collection object can be found in the registration number for items in the photography collection. Letters are used to indicate media format. Equally, filenames are used by the imaging team to indicate when an access master has been modified. There is also a filename convention which allows information to be automatically ingested into the media asset record.¹⁸⁰

Documentation of choices

Both Ward and Lillico spoke about decisions that were made during the digitisation process which affected the digital asset either visually or semantically. These were often referred to as interpretation. When prompted around how these decisions, or the process of making them, was captured or documented it was clear that in some areas this was happening, in others it occurred in a less obvious way, and occasionally these decisions were not documented at all. Because of the focus on the collection object, and the way in which descriptive elements are focused on this record, information about the digital media asset seems to be more prescribed and basic. Information about decisions made relating to background choices, angles, or lighting is not captured routinely. However, this information could be considered to incidentally represent the relationship by implying that the digitisation process has taken place.

¹⁸⁰ Kingston, 'KE EMu Media Assets Manual, General Staff Version', 8.

One thing that was mentioned in the interview with Lillico was where the appropriate place for these decisions to be recorded would be? While there was good awareness that these types of decisions are being made it was less clear how they were recorded by those not directly involved in the decision-making process. For Ward, who is involved in many of the discussions around digitisation decisions, there is more clarity around how these are recorded. This is so that, as a team, they have records so when they come back to imaging a collection, perhaps after a logistical delay, they can be clear about the choices made and why. It might also include specifics around a lighting set-up so it can be recreated. This is stored internally to the team before periodically being uploaded to a shared document on Pou Mataaho, where documentation for procedures and similar are stored for the institution. This is over and above the tracking documents they create for the ACDP, which also include notes around any issues that may have arisen.

While image technicians, such as Ward, do not add information to the collection records about objects they image, they can pass on this information to other staff who can then contribute this to the record. A recent example of this where a moiré was seen on a colour negative they were imaging and they were able to zoom right in and see and identify the process based on this. This was then added to the record by a conservator.

Visual metadata

Some digital imaging decisions are captured within the visual. A prominent example of this is an early decision to crop to the outside of the negative. This decision was made to prove they were showing everything that was on the negative but also allows the materiality of the negative to be seen. There can also be things written on those edges. The objects themselves, not just their content, becomes a history item documenting how photography is made. Another decision that was documented visually was whether blank pages in a volume are photographed. Lillico noted that discussions around this would depend on the nature of the volume, and not just create a solution that is applied universally. In her example, a photograph of an A4 sheet of paper, explaining that there were blank pages which had not been photographed, was

used. Equally, however, information in the collection record may include that there are blank pages.

There is another piece of metadata that gets recorded visually when photographing works on paper, the image of the target. Works on paper being digitised have a standard 45 degrees lighting setup and the evenness of the light is carefully measured. They also shoot a target. The image of the target is attached to the group of objects photographed. Ward notes, “[It’s an] interesting bit of metadata that we generate. So, it’s a photograph to refer to other photographs”. Other multimedia, such as images of paper records, can also be added to records which may provide more information about the collection object. This will often be done by collection managers or conservators. Ward provided an example she had come across recently where photographs of old enclosures were included on a negative’s record after they were rehoused.

Images of the storage of the items may also be captured as part of the ACDP, such as with an image of a whole drawer of nested pounamu. It was originally captured as a way to check records because you can zoom right in and see the registration numbers. It will be ingested into EMu and attached to the storage record and may be useful later for knowing what was in the drawer at that date.

Other decisions can be inferred through the digital image, such as the background colour or that a record for a negative has a positive image attached. Equally, the decision to capture proof sheets, rather than individual images, in order to follow the registration process and current record set-up, can be seen in the image displayed for the record.

More than one version of an image may also be available publicly which can give a sense of the decisions that were made. For colour negatives, the preservation master is created from the light table, white balanced. Then, during inversion for the access master, a decision is made about correcting any colour fade or shift from

degradation. This decision is influenced by many factors but is quite subjective and interpretive. There is a formula that is followed to keep it more of a neutral inversion, but it does require tweaks due to different film bases. When this happens, there are two access masters, one as a straight inversion and the other, with the tweaks, is called a modified master. For internal users, this modification is even clearer as it is captured in the filename with a code.¹⁸¹

Location history: incidental metadata

Metadata can also represent the relationship between a digital surrogate and the collection object in an incidental way. This is clear when looking at the location history for an item. This is updated by a collection manager as an item is moved to the imaging studio. The process of this movement changes the location and also records the reason for the change in location. For an item coming from photography, this may mean many location changes as objects in the cool store are boxed according to the group being imaged, moved into the slightly warmer area to acclimatise, and then into another area to warm up slightly more, before being imaged and this process repeated backwards. Something like this allows for a collection record without an image in the system to be understood to have gone through this process. This has already been used to pick up on something that should have been digitised, but apparently was not.

I just had an inquiry from the curator saying ‘how come these 10 haven’t been done?’ and then looking back and seeing well they were delivered to the photography studio but they’ve obviously been skipped, so they came out and went back in, so we’ll have to redo those ones. (Kirsty Lillico)

This indicates that this piece of metadata does then end up representing this relationship between the existence of the digital surrogate and the collection object.

Summary

At the outset of the research, I aimed to find out how metadata was used at Te Papa to represent the relationship between a digital surrogate and source material. While

¹⁸¹ Dionne Ward (Imaging Technician, Museum of New Zealand Te Papa Tongarewa), interview with Laura Jamieson, October 2019.

ultimately it is clear the relationship is primarily represented in the system architecture, metadata does perform this role in some ways. This research shows relationships are considered important and so are captured in the metadata primarily through the use of the subjects and associations fields, which capture relationships with other collection objects and to general people, places, and events. This is recorded manually which highlights its importance and supports discoverability of the collection items. For digital surrogates, basic pieces of metadata, such as registration numbers, filenames, and media asset record names, represent the relationship back to the physical object. Other metadata, such as location history or the documentation of choices made during digital imaging has a primary purpose unrelated to this relationship. However, it can be used to incidentally represent the relationship if need be. This type of metadata may be more vulnerable as its primary function is not representing the relationship, and its loss could lead to disconnection, as there are no metadata fields specifically for this relationship.

Conclusion

Digital surrogates are an increasing part of museum collections, as digitisation and digital imaging projects continue to expand and increase. Public expectation that collections are online and accessible also encourages the creation of digital surrogates and their use as part of collection management. These digital surrogates have an inherent relationship with their source material which provides context to both the collection object and the new digital asset. This context relies on an awareness of the relationship and the connection between them being represented in some way. The importance of the relationship is touched on in the museums literature, however, the focus is on the possible implications of disconnection and not on how the collections are currently connected or disconnected in museum practice.

The research questions aimed to respond to this gap in the literature by providing a strong base for further research into the long-term implications of the current state of collection management as it relates to the relationship between digital surrogates and source material. To do this the questions sought to understand the current situation at a single institution, and explore how the relationship and connection are represented. The primary research question of this thesis was, ‘in what ways is the relationship between physical collection items and their digital surrogates represented using metadata at Te Papa?’. In order to answer the primary research question, I looked to answer four secondary questions which covered various aspects of collection management.

By asking how metadata is used to represent the relationship between physical objects and their digital surrogates I hoped to better understand how metadata is used in this regard at Te Papa and to get a sense of the current state. This type of information is important as it provides and maintains context for items and keeps these relationships as part of the holistic care of these objects and the digital surrogates created from them. These findings are not transferable as other institutions, both in New Zealand and overseas, are unlikely to have similar features or particulars.

This research only sought to get current state information for a single site. This information could be built upon long term to better understand changes over time and the possible implications of these changes at the same institution. Equally, comparative case studies could also draw out different information. This would be particularly useful to understand whether the domain difference that I found in my research is also found in other museums. Due to the constrained and small-scale nature of my research my interviewees' specialities focused mainly on photography collections, expanding out to a general humanities focus. Research which talks to others in humanities, art, or natural history collections, particularly curators and collection managers would build on this research within Te Papa and expand its focus.

In conducting this research, I set out to understand how metadata is used to represent the relationship between a physical object and its digital surrogate at Te Papa. In asking this question, data was revealed which addressed and highlighted another predominant issue regarding digital surrogates. This is that different worldviews impact the perception of digital surrogates and the relationship they have with their source material. These different worldviews underpin the language, the institutional focus, and how digitisation is prioritised, and these all feed into the relationships and how they are represented. This means that at Te Papa, working in the museum domain, the system architecture drives and represents the relationship. This system was chosen and is used in a way that is influenced by this domain worldview. The differing world views and the systems architecture being the primary way it represented can then lead to the relationship being vulnerable. Metadata within this structure is then not primarily used to represent the relationship between the collection object and digital surrogate, however, there is metadata which can be used to represent the relationship, but this is not their primary purpose.

My literature review worked across domains to get a good base understanding of digital collections, digitisation, metadata, and relationships between collections and objects. Libraries and archives literature were used to supplement ideas found in the

museums literature but also to fill gaps. While these types of literature have differences, they all work together to form a base of literature for all types of cultural and collecting institutions. It also seems that ideas which are fundamental and accepted in libraries and archives literature are not always thought of the same way in museums literature, or in museum practice. The gap between theory and practice, and whether the literature has an impact on everyday work is perhaps seen in the larger domain distinction I encountered in my case study, more than when compared with the literature domains.

Certainly, language and its usage and definitions were a key part of this idea early in my research, but it is clear that despite these differences, the fundamental ideas and principles were understood and worked with. Instead, the way that my interviewees approached the ideas of access and preservation seemed to indicate a distinct domain difference. While the ideas of access and preservation sit across all types of literature, and my interviewees spoke about them in similar ways, the two concepts seemed to be spoken about from a different perspective or worldview.¹⁸² The different approach to using digital surrogates for preservation was more expected, due to the more three-dimensional format of collection objects at Te Papa, compared to a library or archive.¹⁸³ As those I spoke to were more focused on photography collections there was some crossover in thinking with the literature from libraries and archives which also deal with photography. Access, however, contained a larger distinction.

Ultimately projects involving digitising material often have access as a primary driver. The types of access, and whether this is successful, has started to be explored by the literature, and the way that things can be preserved in digital format is related

¹⁸² Ciurea and Filip, 'New Researches on the Role of Virtual Exhibitions in Digitization, Preservation and Valorization of Cultural Heritage', 31; Conway, 'Digital Transformations and the Archival Nature of Surrogates', 54–55; Henning, 'New Media', 306, 309; Jardine, 'Reflections on the Preservation of Recent Scientific Heritage in Dispersed University Collections', 741; Reynold, 'The Digital Initiative in Archives', 19, 27–29; Robinson and Tanner, 'Higher Education Digitisation Service: access in the future preserving the past -- the UK perspective', 66.

¹⁸³ Helfrich, 'Questions of Authenticity: Challenges in Archiving Born-Digital Design Records', 27.

to this.¹⁸⁴ My interviewees also spoke about access in a similar way, citing it as one of the primary drivers for digitisation and a positive outcome of the ACDP. However, the way they spoke about access was different than I expected. For libraries and archives, there exists a fundamental level of access for most public collections.¹⁸⁵ Reading rooms and the expectation of a researcher, or another member of the public, coming in to access materials freely within a purpose-built space is common and expected. Digitisation is spoken about in the literature as the next step in providing access; expanding this access out to those who cannot physically come into the collection space.¹⁸⁶ In contrast, the museums literature focuses on providing access to collections, possibly for the first time.¹⁸⁷ The primary method of access in museums is in exhibitions, where this access is mediated through display choices, exhibitions design, and security and conservation concerns.¹⁸⁸ Some hands-on research may occur but the open invitation to research collection objects in a room which has regular hours is not thought of in the same way. For museums, and this was reflected in the way my interviewees spoke about access, digitisation will provide access to collection objects, possibly for the first time.

Libraries, archives, and museums all recognise the ability for digitising collections to uncover and illuminate parts of their collections which are less well-used, understood, or known about.¹⁸⁹ For museums this is much more distinct. Collection objects may be described in the online catalogue, but the ability for someone to find an object and recognise its relationship to their own research may be limited. Equally the types of collections influence this. Without digitisation, descriptions of objects may seem very similar and a researcher may need mediated access to those items

¹⁸⁴ Besser, 'Digital Longevity', 164–76; Jardine, 'Reflections on the Preservation of Recent Scientific Heritage in Dispersed University Collections', 740–741; Mani, 'Digitisation: Preservation and Challenges', 71; Sabharwal, 'Information Architecture and Hypertextuality: Concerns for Digital Curation', 72.

¹⁸⁵ Conway, 'Digital Transformations and the Archival Nature of Surrogates', 55.

¹⁸⁶ Ciurea and Filip, 'New Researches on the Role of Virtual Exhibitions', 26; Conway, 'Digital Transformations and the Archival Nature of Surrogates', 55; Jardine, 'Reflections on the Preservation of Recent Scientific Heritage in Dispersed University Collections', 740–741.

¹⁸⁷ Henning, 'New Media', 309.

¹⁸⁸ Parry, *Recoding the Museum: Digital Heritage and the Technologies of Change*, 76–77.

¹⁸⁹ Ciurea and Filip, 'New Researches on the Role of Virtual Exhibitions', 26, 30–31; Henning, 'New Media', 306, 309; Jardine, 'Reflections on the Preservation of Recent Scientific Heritage in Dispersed University Collections', 740–741; Reynold, 'The Digital Initiative in Archives', 19, 27–29, 32, 99, 103.

through a curator to ascertain the usefulness to their research.¹⁹⁰ Whereas items in libraries and archives may be distinct and unique or able to be partially or fully transcribed, museum collection objects can be more difficult to describe without the visual representation.

Access is then primarily provided through the digitised images. While this can also be true for libraries and archives, what differs is that the museums literature tends to stop at that point. For my interviewees, who were focused on photography, the digital image functions in place of the collection object, the image itself is the object regardless of format. There is not the expectation that is seen in the libraries and archives literature that it may change access to the object physically, either by increasing or decreasing demand. The idea that physical access may be locked down, with the digitised output as the primary access point, as can happen in libraries and archives space, is not considered in museums, possibly because this level of access does not already exist. More research into this area would be useful to understand the different levels of access and ways of thinking about it across museums, particularly in comparison to libraries and archives. This comparison could perhaps show some of the underlying differences in how these types of institutions function and also collection objects. Equally, work on the successfulness of this digital access could also extend across from libraries and archives literature into the museum domain, and explore whether these different ideas about access impact on the eventual public access in practice.

In asking what metadata was used to represent the relationship between a digital surrogate and the source material I ultimately found that this is not a central concern or focus, as the relationship primarily manifested through the system architecture. However, there are types of metadata which do represent the relationship, albeit incidentally and not as their primary function. These fields, such as the location history, or external documentation about digitisation choices, which contain information which relates back to the record, are created and used for other purposes. Equally, visual metadata that is created may show the current storage or information

¹⁹⁰ Henning, 'New Media', 309.

about the digitisation process is primarily captured to assist the process. These types of metadata may then be vulnerable to being replaced, disconnected, or removed as they are no longer needed for their primary function.

If metadata is not created for the primary purpose of representing this relationship, but other metadata achieves this incidentally, or outside of the collection record, could this connection be vulnerable? In asking my interviewees about the relationship between a digital surrogate and the source material it was apparent that their focus is not on this, and they see this relationship sitting implicitly within the system architecture. Where metadata does represent this relationship, it is usually a secondary function of that information and not actively thought of as useful for this. This, as expressed earlier, may be a reflection of the different ways of thinking between the domains of museums, libraries, and archives, with the museum more focused on a holistic view of the relationships, rather than embedding this information into the record.

Digital surrogates, in this case study, were not thought of as independent collection items, rather part of the overall collection record and information. This limits the metadata used to represent the relationship as well, as it is assumed it is not needed and the system does not require it. Despite this, the system does treat these digital assets as individual items and stores appropriate information, which alleviates some of the concerns around vulnerability. Where I have found metadata representing the relationship, however, it has usually been a secondary focus of that information, such as the use of location history to track movement to and from the digitisation studio. This may make this metadata vulnerable to system migration. Just as information may have been stored in a card catalogue system and then not transferred to an electronic system, moving between systems can require rethinking the way information is stored. Data costs money, space, and other resources to store. While hugely important to collections, it does not seem out of the realms of possibility that a pragmatic approach to this may involve looking at the information that is stored and trying to consolidate it. Location history is unlikely to be a high priority piece of information to be retained.

This sort of vulnerability may be unlikely in a large and well-resourced museum like Te Papa, but it does highlight ongoing concerns in the literature about disconnection and loss of context that can happen without adequate metadata. Jones' discussion of field diaries in museum collections is an example of this disconnection within an institution, where the inherent relationship between the field diaries and objects collected on those expeditions would have been clear for both the institution and external researchers for a long time.¹⁹¹ However, due to being treated as part of different collections and with inadequate metadata to represent relationships, these relationships broke down, particularly from an external point of view.

The implications for the lack of metadata specifically designed to represent the relationship between the digital surrogate and source material is outside the scope of my research, as is whether the use of the system architecture or more incidental metadata would be sufficient long term. However, speaking to Lillico, McCredie, Moorhead, and Ward, and reading some of the documentation Te Papa has around the creation of metadata, indicates to me that while this area is not something they may have actively thought about, it influences other decisions that they make, especially in the way that Te Papa's systems are set up, and the current digitisation focus. This more holistic way of thinking, one which may be considered less database minded, may mean the lack of focus on the relationship between digital surrogates and the source material in the metadata does not lead to a vulnerable relationship, however, I do think it is an interesting thing to consider as institutions increase their digital collections.

Having an understanding of how the relationship between physical objects and digital surrogates is represented at Te Papa provides a base for further research in this area to better understand differences across domains, as well as the implications for particular practices. This research contributes to the literature by uncovering differences across domains, related to worldview, language, and perspectives, which

¹⁹¹ Jones, 'Documenting Artefacts and Archives in the Relational Museum', 165–166.

reflect the range of domains which were consulted in the literature. Better understanding these domains distinctions, as well as their similarities or overlaps, could support research across the cultural and collecting fields, and create positive changes in practice to better understand the different mandates institutions have, and also how they can work together.

This research also provides information and analysis of the use of the system architecture to represent the relationship and areas of metadata which could also be used in this way if required. Further research at Te Papa could build on this, as well as explore any changes to this practice over time, helping to better understand the implications and possible vulnerabilities. Future research into the fundamental relationship between digital surrogates and source material would support collection management practices which provide context and information to collection objects, as well as contribute to best practice for ongoing and future digitisation projects.

Digital surrogates form a key part of digital collections, and their relationship to the source material is important in order to provide context and information. This research looked at the current state at a single site, and found that the way that this relationship is represented is influenced by the museum context. While museums could benefit from knowledge from the library and archive domains in thinking about digital collections and their relationships to physical collection objects, care should be taking in assuming that those processes and ideas could directly translate into the museum context. Differences in language, ways of working, and collecting mandates need to be taken into consideration to ensure that practical solutions and processes suit the museum collections and the museum context. Ultimately, the importance of relationships in collection management is seen across the domains. This research shows that the way that we represent a relationship is equally significant.

Bibliography

Anderson, Maxwell L. 'Electronic Information and Digitization: Preservation and Security Challenges'. In *The Strategic Stewardship of Cultural Resources: To Preserve and Protect*, edited by Andrea T. Merrill, 163–68. Florence, United Kingdom: Routledge, 2003.

Arroyo-Ramírez, Elvia, Kelly Bolding, Faith Charlton, and Allison Hughes. “‘Tell Us about Your Digital Archives Workstation’: A Survey and Case Study”. *Journal of Contemporary Archival Studies* 5, no. Article 16 (2018).

Bearman, David. 'Possible Contributions of the Reference Model of Metadata Required for Evidence to a Reference Model of Metadata Required for Image Description'. *Archives and Museum Informatics* 10, no. 3 (December 1996): 295–302. <https://doi.org/10.1007/BF02802377>.

Beaudoin, Joan E. 'A Framework for Contextual Metadata Used in the Digital Preservation of Cultural Objects'. *D-Lib Magazine* 18, no. 11/12 (November 2012): 1–26. <https://doi.org/10.1045/november2012-beaudoin2>.

Beaudoin, Joan E. 'Context and Its Role in the Digital Preservation of Cultural Objects'. *D-Lib Magazine* 18, no. 11/12 (November 2012): 1–27. <https://doi.org/10.1045/november2012-beaudoin1>.

Besser, Howard. 'Digital Longevity'. In *Handbook For Digital Projects: A Management Tool for Preservation and Access*, edited by Maxine K. Sitts, 164–76, First Edition. Massachusetts, USA: Northeast Document Conservation Center, 2000.

Blackaby, Jim, and Beth Sandore. 'Building Integrated Museum Information Retrieval Systems: Practical Approaches to Data Organization and Access'. *Archives and Museum Informatics* 11 (1997): 117–46.
<https://doi.org/10.1023/A:1009096811050>.

Bremer-Laamanen, Majlis. 'Digitisation for Access to Preserved Documents'. *Liber Quarterly: The Journal of European Research Libraries* 13, no. 2 (2003): 137–145.
<https://doi.org/10.18352/lq.7726>.

Brocks, Holger, Alfred Kranstedt, Gerald Jäschke, and Matthias Hemmje. 'Modeling Context for Digital Preservation'. In *Smart Information and Knowledge Management: Advances, Challenges, and Critical Issues*, edited by Edward Szczerbicki and Ngoc Thanh Nguyen, 197–226. Studies in Computational Intelligence. Berlin, Heidelberg: Springer Berlin Heidelberg, 2010.
https://doi.org/10.1007/978-3-642-04584-4_9.

Bryman, Alan. *Social Research Methods*. Fifth edition. Oxford; New York: Oxford University Press, 2016.

Burgess, Matthew. 'Digital preservation at the point of acquisition: Collecting born-digital photographs'. State Library of NSW, 11 February 2019. Accessed 4 April 2019. <https://www.sl.nsw.gov.au/blogs/collecting-born-digital-photographs>.
(<https://web.archive.org/web/20190404040250/https://www.sl.nsw.gov.au/blogs/collecting-born-digital-photographs>).

Cameron, Fiona. 'Museum Collections, Documentation, and Shifting Knowledge Paradigms'. In *Reinventing the Museum: The Evolving Conversation on the Paradigm Shift*, edited by Gail Anderson, 2nd edition, 223–38. Lanham, Md.: AltaMira Press, 2012.

Campbell, Danielle. 'Making matches: Reconnecting Leslie Adkin photographs'. Te Papa's Blog, 31 January 2020. Accessed 3 February 2020.

<https://blog.tepapa.govt.nz/2020/01/31/making-matches-reconnecting-leslie-adkin-photographs/>.

(<https://web.archive.org/web/20200203030119/https://blog.tepapa.govt.nz/2020/01/31/making-matches-reconnecting-leslie-adkin-photographs/>).

Ciurea, Cristian, and Florin Filip. 'New Researches on the Role of Virtual Exhibitions in Digitization, Preservation and Valorization of Cultural Heritage'.

Informatica Economica 20, no. 4 (2016): 26–33.

<https://doi.org/10.12948/issn14531305/20.4.2016.03>.

Conway, Paul. 'Archival Quality and Long-Term Preservation: A Research Framework for Validating the Usefulness of Digital Surrogates'. *Archival Science* 11, no. 3 (1 November 2011): 293–309. <https://doi.org/10.1007/s10502-011-9155-0>.

Conway, Paul. 'Digital Transformations and the Archival Nature of Surrogates'. *Archival Science* 15, no. 1 (2014): 51–69. <https://doi.org/10.1007/s10502-014-9219-z>.

Conway, Paul. 'Digitizing Preservation'. *Library Journal* 119, no. 2 (1994): 42–45.

Conway, Paul. 'Rationale for Digitization and Preservation'. In *Museums in a Digital Age*, edited by Ross Parry, 365–90. Florence, United Kingdom: Routledge, 2010.

Crow, Jeff, Luis Francisco-Revilla, April Norris, Shilpa Shukla, and Ciaran B. Trace. 'A Unique Arrangement: Organizing Collections for Digital Libraries, Archives, and Repositories'. In *Theory and Practice of Digital Libraries*, edited by Panayiotis

Zaphiris, George Buchanan, Edie Rasmussen, and Fernando Loizides, 335–44. *Lecture Notes in Computer Science*. Springer Berlin Heidelberg, 2012.

Deegan, Marilyn, and Simon Tanner. *Digital Futures: Strategies for the Information Age*. London: Library Association, 2002.

Denscombe, Martyn. *The Good Research Guide: For Small-Scale Social Research Projects*. England: McGraw-Hill/Open University Press, 2014.

Duranti, Luciana. ‘The Archival Bond’. *Archives and Museum Informatics* 11, no. 3 (1 September 1997): 213–18. <https://doi.org/10.1023/A:1009025127463>.

Duval, Erik, Wayne Hodgins, Stuart Sutton, and Stuart L. Weibel. ‘Metadata Principles and Practicalities’. *D-Lib Magazine* 8, no. 4 (April 2002). <https://doi.org/10.1045/april2002-weibel>.

Federal Agencies Digital Guidelines Initiative. “Glossary; Term: Digital Imaging”. Accessed 29 January 2020. <http://www.digitizationguidelines.gov/term.php?term=digitalimaging>. (<https://web.archive.org/web/20200129010531/http://www.digitizationguidelines.gov/term.php?term=digitalimaging>).

Field, Jeffrey M. ‘Building a National Preservation Program: National Endowment for the Humanities’. In *The Strategic Stewardship of Cultural Resources: To Preserve and Protect*, edited by Andrea T. Merrill, 59–66. Florence, United Kingdom: Routledge, 2003.

Findlay, Peter. 'The UK Medical Heritage Library and the Relationship Between Print and the Digital'. *Journal of Victorian Culture* 23, no. 2 (27 April 2018): 231–37. <https://doi.org/10.1093/jvcult/vcy021>.

Fleischhauer, Carl. 'Preservation, Security, and Digital Content'. In *The Strategic Stewardship of Cultural Resources: To Preserve and Protect*, edited by Andrea T. Merrill, 141–47. Florence, United Kingdom: Routledge, 2003.

Goss, Suzanne Marie. 'Developing Digital Capability: What Archivists Can Learn from the GLAM Sector'. Victoria University of Wellington, 2017.

Gould, Sara, Marie-Therese Varlamoff, Richard Ebdon, International Federation of Library Associations, and UNESCO. *IFLA/UNESCO Survey on Digitisation and Preservation*. International Preservation Issues; No. 2. Wetherby, UK; Paris, France: IFLA-UAP; IFLA-PAC, 1999.

Groskopf, Christopher. 'Museums Are Keeping a Ton of the World's Most Famous Art Locked Away in Storage'. Quartz, 21 January 2016. Accessed 26 February 2020. <https://qz.com/583354/why-is-so-much-of-the-worlds-great-art-in-storage/>. (<https://web.archive.org/web/20200225230057/https://qz.com/583354/why-is-so-much-of-the-worlds-great-art-in-storage/>).

Guest, Greg, Kathleen MacQueen, and Emily Namey. *Applied Thematic Analysis*. California: SAGE Publications, Inc., 2012. <https://doi.org/10.4135/9781483384436>.

Helfrich, Kurt G. F. 'Questions of Authenticity: Challenges in Archiving Born-Digital Design Records'. *Art Libraries Journal* 35, no. 3 (2010): 23–29. <https://doi.org/10.1017/S0307472200016503>.

Henning, Michelle. 'New Media'. In *Companion to Museum Studies*, edited by Sharon Macdonald, 302–318. Wiley, 2008.

Higgins, Sarah. 'Digital Curation: The Emergence of a New Discipline'. *International Journal of Digital Curation* 6, no. 2 (2011): 78–88.
<https://doi.org/10.2218/ijdc.v6i2.191>.

Jardine, Nicholas. 'Reflections on the Preservation of Recent Scientific Heritage in Dispersed University Collections'. *Studies in History and Philosophy of Science* 44, no. 4 (2013): 735–743. <https://doi.org/10.1016/j.shpsa.2013.07.009>.

Jones, Michael Alastair. 'Documenting Artefacts and Archives in the Relational Museum'. PhD, The University of Melbourne, 2018.

Jones, Mike. 'On Digital Archives'. *Context Junky* (blog), 21 November 2012. Accessed 4 December 2017.
<http://www.mikejonesonline.com/contextjunky/2012/11/21/on-digital-archives/>.
(<https://web.archive.org/web/20171204170350/http://www.mikejonesonline.com/contextjunky/2012/11/21/on-digital-archives/>).

Jones, Mike. 'Physical Office, Digital Outhouse'. *Context Junky* (blog), 28 July 2015. Accessed 28 November 2019.
<http://www.mikejonesonline.com/contextjunky/2015/07/28/physical-office-digital-outhouse/>.
(<https://web.archive.org/web/20191128174014/http://www.mikejonesonline.com/contextjunky/2015/07/28/physical-office-digital-outhouse/>).

King, Nigel, and Christine Horrocks. *Interviews in Qualitative Research*. Los Angeles: SAGE, 2010.

Knell, Simon J. 'The Shape of Things to Come' Museums in the Technological Landscape'. In *Museums in a Digital Age*, edited by Ross Parry, 435–53. Florence, United Kingdom: Routledge, 2010.

Knell, Simon J. *Museums and the Future of Collecting*. Florence: Routledge, 2004.

Lee, Christopher A. 'A Framework for Contextual Information in Digital Collections'. *Journal of Documentation* 67, no. 1 (2011): 95–143.

Mani, N. 'Digitisation: Preservation and Challenges'. *DESIDOC Journal of Library & Information Technology* 29, no. 1 (2009): 69–71.
<https://doi.org/10.14429/djlit.29.233>.

Marshall, MN. 'The Key Informant Technique'. *Family Practice* 13, no. 1 (1996): 92–97.

McCarthy, Conal. *Te Papa: Reinventing New Zealand's National Museum, 1998-2018*. Wellington, New Zealand: Te Papa Press, 2018.

Mitcham, Jenny. 'Emulation for Preservation - Is It for Me?' *Digital Archiving at the University of York* (blog), 23 June 2017. Accessed 11 February 2020. https://digital-archiving.blogspot.com/2017/06/emulation-for-preservation-is-it-for-me_27.html.
(https://web.archive.org/web/20200211001414/https://digital-archiving.blogspot.com/2017/06/emulation-for-preservation-is-it-for-me_27.html).

Moorhead, Fiona. "'Good-bye, moon": completing the Leslie Adkin diary project'. Te Papa's Blog, 21 December 2018. Accessed 3 February 2020.

<https://blog.tepapa.govt.nz/2018/12/21/completing-adkin-diary-project/>.
(<https://web.archive.org/web/20200203025903/https://blog.tepapa.govt.nz/2018/12/21/completing-adkin-diary-project/>).

Moorhead, Fiona. 'Te Papa's New Collections Site Launches – Get Hunting'. Te Papa's Blog, 14 September 2017. Accessed 7 January 2020.
<https://blog.tepapa.govt.nz/2017/09/15/te-papas-new-collections-site-launches-get-hunting/>.
(<https://web.archive.org/web/20200106234456/https://blog.tepapa.govt.nz/2017/09/15/te-papas-new-collections-site-launches-get-hunting/>).

Moran, Jessica. 'Born Digital in New Zealand: Report of Survey Results'. Wellington, New Zealand: Alexander Turnbull Library, National Library of New Zealand, 2017. <https://natlib.govt.nz/files/reports/research-borndigital2017-report.pdf>.

Moran, Jessica. 'The Days of Our (Digital) Lives'. *Blog / National Library of New Zealand* (blog), 3 October 2018. Accessed 24 January 2019.
<https://natlib.govt.nz/blog/posts/the-days-of-our-digital-lives>.
(<https://web.archive.org/web/20190124104558/https://natlib.govt.nz/blog/posts/the-days-of-our-digital-lives>).

Moriarty, Adam. 'A Crisis of Capacity: How Can Museums Use Machine Learning, the Gig Economy and the Power of the Crowd to Tackle Our Backlogs – MW19 | Boston'. Boston, 2019. Accessed 20 May 2019. <https://mw19.mwconf.org/paper/a-crisis-of-capacity-how-can-museums-use-machine-learning-the-gig-economy-and-the-power-of-the-crowd-to-tackle-our-backlogs/>.
(<https://web.archive.org/web/20190520022928/https://mw19.mwconf.org/paper/a-crisis-of-capacity-how-can-museums-use-machine-learning-the-gig-economy-and-the-power-of-the-crowd-to-tackle-our-backlogs/>).

Museum of New Zealand Te Papa Tongarewa, Wellington, New Zealand. 'About Collections Online,' 16 August 2017. Accessed 7 January 2020.

<https://www.tepapa.govt.nz/discover-collections/collections-online/about-collections-online>.

(<https://web.archive.org/web/20200107000201/https://www.tepapa.govt.nz/discover-collections/collections-online/about-collections-online>).

Museum of New Zealand Te Papa Tongarewa, Wellington, New Zealand. 'National Services Te Paerangi,' 9 February 2016. Accessed 26 February 2020,

<https://www.tepapa.govt.nz/learn/for-museums-and-galleries/national-services-te-paerangi>.

(<https://web.archive.org/web/20200226010829/https://www.tepapa.govt.nz/learn/for-museums-and-galleries/national-services-te-paerangi>).

Museum of New Zealand Te Papa Tongarewa, Wellington, New Zealand. 'New Collections Online Launches,' 15 September 2017. Accessed 7 January 2020.

<https://www.tepapa.govt.nz/about/news/new-collections-online-launches>.

(<https://web.archive.org/web/20200106235451/https://www.tepapa.govt.nz/about/news/new-collections-online-launches>).

Museum of New Zealand Te Papa Tongarewa, Wellington, New Zealand. 'Te Papa's Vision and Future,' 10 February 2016. Accessed 9 December 2019.

<https://www.tepapa.govt.nz/about/what-we-do/te-papas-vision-and-future>.

(<https://web.archive.org/web/20191209005755/https://www.tepapa.govt.nz/about/what-we-do/te-papas-vision-and-future>).

National Digital Forum. 'Past conferences'. Accessed 26 February 2020.

<http://www.ndf.org.nz/past-conferences-1>.

(<https://web.archive.org/web/20200226014129/http://www.ndf.org.nz/past-conferences-1>).

Note, Margot. *Managing Image Collections*. Chandos Information Professional Series. Oxford; Cambridge; New Delhi: Chandos Publishing, 2011.

Orr, Renee. 'Digitisation and Workplace Learning: An Exploratory Study'. Victoria University of Wellington, 2006.

Oxford University Press. "Metadata," OED Online. Accessed 29 January 2020.
<https://www.oed.com/view/Entry/117150?redirectedFrom=metadata#eid37413841>.

Padfield, Tim. 'Preserving and Accessing Our Cultural Heritage - Issues for Cultural Sector Institutions: Archives, Libraries, Museums and Galleries'. In *Copyright and Cultural Heritage: Preservation and Access to Works in a Digital World*, edited by Estelle Derclaye, Chapter 9. Cheltenham, UK; Northampton, MA, USA: Edward Elgar, 2010.

Parry, Ross. *Museums in a Digital Age*. Florence: Routledge, 2010.

Parry, Ross. *Recoding the Museum: Digital Heritage and the Technologies of Change*. London, United Kingdom: Routledge, 2007.

Payne, Geoff, and Judy Payne. 'Key Informants'. In *Key Concepts in Social Research*, 135–38. London: SAGE Publications, Ltd, 2011.
<https://doi.org/10.4135/9781849209397>.

Ray, Joyce. 'The Rise of Digital Curation and Cyberinfrastructure'. *Library Hi Tech* 30, no. 4 (2012): 604–622. <https://doi.org/10.1108/07378831211285086>.

Redwine, Gabriela, Megan Barnard, Kate Donovan, Erika Farr, Michael Forstrom, Will Hansen, Jeremy Leighton John, Nancy Kuhl, Seth Shaw, and Susan Thomas. 'Born Digital: Guidance for Donors, Dealers, and Archival Repositories'. Washington, DC: Council on Library and Information Resources, October 2013.

Reynold, Kate Louise Musker. 'The Digital Initiative in Archives: A Study on the Selection of Archives for Digitisation in New Zealand'. Research Paper MLIS, Victoria University of Wellington, 2006.

Robinson, Brian, and Simon Tanner. 'Higher Education Digitisation Service: access in the future preserving the past -- the UK perspective'. *Bibliothek Forschung und Praxis* 23, no. 3 (1999): 66–69. <https://doi.org/10.1515/bfup.1999.23.1.66>.

Sabharwal, Arjun. 'Information Architecture and Hypertextuality: Concerns for Digital Curation'. In *Digital Curation in the Digital Humanities*, edited by Arjun Sabharwal, 69–93. Chandos Publishing, 2015. <https://doi.org/10.1016/B978-0-08-100143-1.00004-0>.

Shein, Cyndi, and Emily Lapworth. 'Say Yes to Digital Surrogates: Strengthening the Archival Record in the Postcustodial Era'. *Journal of Western Archives*, 7, no. 1 (2016): 52.

Singer, Randy, and Gil Nelson. 'The Impact of Digitization and Digital Data Mobilization on Biodiversity Research and Outreach'. Presented at the SPNHC 2018, 29 August 2018. <https://twitter.com/SiobhanLeachman/status/1034547756189212672>.

Smith, Brett. 'Generalizability in Qualitative Research: Misunderstandings, Opportunities and Recommendations for the Sport and Exercise Sciences'. *Qualitative Research in Sport* 10, no. 1 (October 2018).
<http://dx.doi.org/10.1080/2159676X.2017.1393221>.

Terras, Melissa M. 'The Rise of Digitization'. In *Digitisation Perspectives*, edited by Ruth Rikowski, 3–20. Educational Futures Rethinking Theory and Practice. Rotterdam: SensePublishers, 2011. https://doi.org/10.1007/978-94-6091-299-3_1.

Torremans, Paul. 'Archiving Exceptions: Where Are We and Where Do We Need to Go?' In *Copyright and Cultural Heritage: Preservation and Access to Works in a Digital World*, edited by Estelle Derclaye, Chapter 5. Cheltenham, UK; Northampton, MA, USA: Edward Elgar, 2010.

Trant, J. 'When All You've Got Is "The Real Thing": Museums and Authenticity in the Networked World'. *Archives and Museum Informatics* 12, no. 2 (1 June 1999): 107–25. <https://doi.org/10.1023/A:1009041909517>.

Unsworth, John. 'The Value of Digitization for Libraries and Humanities Scholarship'. The Newberry Library, 2004. Accessed 11 September 2018.
<http://www.people.virginia.edu/~jmu2m/newberry.04.html>.
(<https://web.archive.org/web/20180911015145/http://www.people.virginia.edu/~jmu2m/newberry.04.html>).

Van de Sompel, Herbert, and Patrick Hochstenbach. 'Reference Linking in a Hybrid Library Environment: Part 1: Frameworks for Linking'. *D-Lib Magazine* 5, no. 4 (April 1999): 19. https://doi.org/10.1045/april99-van_de_sompel-pt1.

Wallace, Andrea, and Ronan Deazley. 'Digital Surrogates'. Display At Your Own Risk, 2016. Accessed 5 February 2020. <https://displayatyourownrisk.org/digital-surrogates/>.

(<https://web.archive.org/web/20200205031405/https://displayatyourownrisk.org/digital-surrogates/>).

Watkins, Gareth. 'Rapidly Digitising 10,000 Collection Items'. Te Papa's Blog, 26 February 2018. Accessed 22 November 2019.

<https://blog.tepapa.govt.nz/2018/02/27/a-digital-birthday-treat/>.

(<https://web.archive.org/web/20191122062345/https://blog.tepapa.govt.nz/2018/02/27/a-digital-birthday-treat/>).

Wecker, Menachem. 'Are Museums Being Clear Enough with the Public about What's Real and What's Fake?' *Washington Post*, 27 February 2019. Accessed 3 April 2019.

<https://www.washingtonpost.com/news/magazine/wp/2019/02/27/feature/are-museums-being-clear-enough-with-the-public-about-whats-real-and-whats-fake/>.

(<https://web.archive.org/web/20190403171902/https://www.washingtonpost.com/news/magazine/wp/2019/02/27/feature/are-museums-being-clear-enough-with-the-public-about-whats-real-and-whats-fake/>).

Wellington, Shannon. 'Building GLAMour: Converging Practice between Gallery, Library, Archive and Museum Entities in New Zealand Memory Institutions'. PhD, Victoria University of Wellington, 2013.

Wellington, Shannon, and Gillian Oliver. 'Reviewing the Digital Heritage Landscape: The Intersection of Digital Media and Museum Practice'. In *The International Handbooks of Museum Studies*, edited by Sharon Macdonald and Helen Rees Leahy, 577–98. Oxford, UK: John Wiley & Sons, Ltd, 2013.

<https://doi.org/10.1002/9781118829059.wbihms225>.

Appendices

Appendix one: Interview guide

Numbered points indicate general order with the exception that 4 may come above 2 and 3 if their role, as described in 1, is more focused on metadata. Secondary bullet-points are possible prompts or probes.

1. Background

- Tell me about your role
- How does your role relate to digitisation or digital imaging?
- What part do you play in the digitisation process?

2. Digitisation/digital imaging

- What is the process for deciding an object will be digitised?
- How do you find information about whether something is already digitised?
- How would you define a digital surrogate?
 - (Follow-up) *Define digital surrogate for them?*
- What types of digital surrogates are made of collection items here that you are aware of?
 - (Prompt) Images, 3D scans, models?

3. Role of digital surrogates

- What do you see as the function of digital surrogates?
 - (Prompt) Access/Preservation/Awareness
- Are digital surrogates used in this way here?
- Do you think the use of digital surrogates has changed?
 - (Probe) In what ways?
- Do you think the use of digital surrogates will change?
 - (Probe) In what ways?
- Are digital surrogates ever thought of as distinct collection items?
 - (Prompt) Separate collection objects, not just part of collection record

- (Probe) How do you feel about that?

4. Metadata creation

- Tell me about your role, and how it relates to metadata
 - (Prompt) What role do you play in creating, editing or recording metadata?
- Are there policies or guidelines about creating metadata?
 - (Probe) What are they?
 - (Probe) Are those written down?
- Do you think metadata creation is consistent across objects?

5. Metadata and digitisation

- How do you tell in the CMS that something has been digitised?
- What metadata is added to physical objects when they are digitised?

6. The relationship between them

- Do you think digital surrogates have a relationship with the object they were created from/of?
- Can you describe what this relationship could be and why it may be important?
- How important do you think the relationship is between a digital surrogate and the object?
- Do you think metadata can represent the relationship that a digital surrogate and a physical object have?
- What do you think is important information about the physical item that should be recorded with the digital surrogate?

7. Metadata representing this

- What metadata relates the digital surrogate to the physical object?
- What metadata relates the physical object to the digital surrogate?

- Can you think of any examples of metadata which refers the digital surrogate to the physical object or vice versa?
- Earlier we discussed metadata policies and guidelines, do you think that the policies account for this relationship?
- Does any information about digital surrogates come from the physical item collection record?

8. Collections online

- How is it clear in collections online that something has been digitised?
- Which metadata fields are shown on collections online?

9. Close

- What do you hope to see happen in this general area moving forward?
- Is there anything else you wanted to share with me?
- Is there anything you wanted to add that we haven't discussed?
- Do you have questions or thoughts about this research?

Appendix two: Codes and themes

Descriptive codes

Name of code	Code boundaries
Access	Discussion of access in general, particularly access to content and digital collections or information
Access (physical)	When discussion of access related to physical access to collection items
CMS info	General information about the CMS, EMu
CMS structure	Information about the structure of the CMS or the records
Digitisation (choices)	Discussion on the choices made in digitisation, this may overlap with 'Interpretation' and 'Priorities' but is a more general category
Digitisation (process)	General discussion of digitisation, particularly the process involved, including image processing
Discoverability	Specific references to the idea of discoverability. May overlap with 'Access'
Interpretation	Specific references to the idea of interpretation in digitisation. Also includes the implications and effect of these choices. May overlap with 'Digitisation (choices)'
Legacy Images	Referring to digital media assets which are older than other digital media assets for the same item, i.e. where that item has been re-digitised.
Location	References to the location of an item, either the metadata which discusses this or where it is held

Metadata	Any references or discussion of metadata, as defined by the research project
Metadata scope	What the interviewee considered metadata to refer to
Preservation	Discussion of preservation of objects, digital assets, or in general
Priorities	Priorities for digitisation or other collection focused work. May overlap with ‘Digitisation (choices)’
Surrogate	Direct or inferred references to digital surrogates, as defined by the research project

Themes

Larger theme	Codes involved
Structure and relationships	CMS structure; Legacy images; CMS info
Definitions, perspectives, language	Metadata scope; Surrogate
Digitisation for access and preservation	Access; Preservation; Priorities; Discoverability; Digitisation (process)
Metadata use	Digitisation (choices); Interpretation; Metadata; Access (physical); Location