

# Archives and Recordkeeping

## An Australian Perspective

Cassie Findlay (Senior Analyst, Information Governance at Gap Inc.)

### Introduction

Today I would like to share with you a few things about recordkeeping and archives in general and about some of the most important techniques for any recordkeeping professional, including archivists.

Yuriko told me that many of you are studying librarianship, so first I want to give you a short description of what the work of an archivist in Australia entails. I hope I am not explaining things to you that you already know! Later on I will do some comparison of the work of librarians and archivists.

I have always put ‘archivist’ in the job section of immigration forms at airports, but in my career I have worked in corporations, governments and as a consultant, in roles that have required me to practise traditional archives management, policy work, digital preservation and most recently, information governance, with a global fashion retailer, based in San Francisco.

This is not uncommon for Australian records and archives people, where we see the work of archivists and records managers as essentially the same. It’s just the contexts we operate in that are different.

Our job is to build and manage systems for records, and to make sure that records are trustworthy and usable, over time.

And a short note about me and my work—at the moment I work for Gap, the global fashion retailer. Gap owns many brands including Banana Republic, Gap and Old Navy. It has stores around the globe—including in Japan—and a huge online business.

In the past I have worked in government and small private sector companies but this is the biggest corporation I’ve worked for, at its headquarters in San Francisco. It’s dynamic and complex, and always changing. I have worked as the corporate archivist and now in the privacy team, helping with ensuring customer data is properly collected, used and shared. I hope I can show you in this talk how the core skills of recordkeeping were essential to me in both roles, and that while in the archive we get to play with the ‘fun’ stuff, like the advertising materials, objects and imagery, both roles are basically about how the company makes and keeps evidence of what it does.

### Archivists and librarians: collaboration or convergence?

But before I go too deep into recordkeeping, I want to briefly touch on some issues of professional identity for archivists and librarians. What do libraries and archives have in common? Archives and libraries both:

- manage and enable access to information resources

- help protect and preserve documentary heritage
- offer online access to an increasing number of their services

However, as you can tell from my talk so far, there are some important differences in the work we do as librarians and archivists—which should not be ignored as budgets shrink, and decisions are made to merge or streamline library and archives services. The table below was compiled by the Australian Society of Archivists<sup>1)</sup> and I think is a useful summary of some of the key differences in what we do.

### Summary of the differences between Archives and Libraries

Archives	Libraries
Specialise in the identification, protection and provision of access to records	Specialise in the acquisition and management of <i>published information</i>
Deal with records in aggregate, often in large volumes, based on the context of their creation	Usually manage items at a more discrete level, individual level
Administer a variety of changing access requirements for records over time, including in law, government or corporate policy, copyright and more	Provide open access to the majority of their holdings, with access restrictions usually based in copyright law
Support accountability through the keeping of good evidence, as well as providing access to records as appropriate	Are primarily concerned with providing easy, equitable access to information
Are expert in creating and managing contextual information on records provenance, and monitoring changes to this over time	Are expert in library cataloguing systems that apply subject based classifications
Make and execute decisions about the destruction of records that must take account of law, societal expectations and other requirements.	Remove items based on library policy and stakeholder engagement

Of course, despite these differences, we can collaborate and work in cross-disciplinary teams, as information management professionals. Indeed, there are many examples of this, especially in the digital preservation world. However we must remember the reasons for our differing approaches, and the outcomes we are working towards. We have distinct professional identities, and separate bodies of knowledge and skills.

Both professions have changed a lot in the last couple of decades, with the digital revolution. For me and my colleagues in Australia, moving from a paper to a digital mindset has been absolutely essential to meeting the challenges of this new world. However there are still core understandings in our professional knowledge that remain the same for all recordkeeping professionals, whether they are working in entirely digital or mixed digital / paper environments. I would like to explain these to you.

## About records

OK, so what do I mean by records? We define these as any information—today, collections of data—that is evidence of a business transaction. So a record is not defined by format, but rather by its purpose—regardless of whether it is paper, digital or an ancient tablet!

Anything that shows us what occurred at a point in time in a particular context can be a record, from a note left on your desk by a colleague, to a set of data that shows you updated your Facebook status. These are all records, it's just that some are managed more carefully, and with more rules, than others. Archival records are simply records that a community, organisation or whole society have applied special rules to, because they are evidence of business or events that the group deems to be important for a variety of different reasons, depending on the group and what matters to them.

Of course records were once all made and kept in paper and other analog forms. Today they are generated in huge quantities in digital form. Think about how rare and precious letters from a century ago can be—and then think about how many messages go through Facebook's servers in a day. It's a very different world!

Records today are made from data—data that shows an event, in context. An event could be an automated one, like the updating of Facebook's databases, it could be a mixture of human action and computer, like the sending of an email, or it could be entirely human, like the writing of the note on paper.

But for all these records, metadata is an essential component.

## About metadata for records

Metadata for records is the information that shows us the context for the event that occurred: whose Facebook status was updated? When? What was the IP address of their computer? Even for paper records there is metadata—it just used to be written or printed on letterheads, file covers and in indexes.

Records can have more or less metadata, depending on what sort of systems they are produced by and maintained in.

Metadata is not a one-off thing; it accumulates through time as records take part in different business transactions—including when they enter the archival domain. Later on I will talk about my work with 'born digital' records as archives—from databases and websites to email inboxes. Analysing metadata and defining new metadata was critical for the success of that work.

Metadata is the essential indicator of context. There are lots of great examples of this. Imagine looking at a series of emails but not being able to see who the sender and recipient was. Or your lecturer checking a record of your submission of an assignment but not being able to see that it was done before the due date.

In archives, metadata is vital for the meaning and usability of all records, but perhaps especially for digital ones. When you work with digital records there are often many thousands or millions of items. We need data points like unique iden-

tifiers and creator/system names to make sense of them. Dates are always important, and if you are working in a records management role, they're essential if you are checking things like whether it is ok to delete a record.

So there is metadata generated as a natural part of business, but as record-keeping professionals, we also add our own metadata to records and sets of records to assist in their management. Access rules, for example, or descriptions of the creating organisation or person.

So clearly, working with metadata is a core skill for recordkeeping professionals, and we, like librarians, have developed standards for the metadata that we need. These help us to implement metadata to make the records we keep more reliable and more usable, through systems design and in business rules. Such standards extend into the archival domain, where we have defined rules for metadata that contextualises and helps us manage archival records—such as archival description standards. I will return to these again shortly.

## The role of archivists and other recordkeeping professionals

We recordkeeping professionals—including archivists—spend a lot of our time thinking about how to create and keep evidence—evidence of business activity in the form of records—whether that business is running a government, providing archival services or selling fashion. Whether I am working as an information governance professional or an archivist, I am actively working to manage records, provide—or restrict—access to them, and ensure they retain authenticity, meaning and usability over time.

To achieve these goals, our core work includes:

- continuing and recurrent analysis of organisational context, business activity, technology and risk;
- defining rules for metadata implementation that helps us contextualise and manage records, over time—including for archival description purposes;
- defining and implementing rules for access and providing access and use of records (or restricting it), and;
- managing continued records retention, systems migrations and deletion.

Depending on where we work, we might also be engaged in:

- copying/digitisation programs;
- managing public access eg under transparency laws, or archival access;
- digital preservation; and
- conservation of physical items.

In the digital world, one of the most important things for recordkeeping professionals of all types to be is also about being proactive, not reactive. How? I will explain some of the techniques that Australian recordkeeping professionals—including archivists—use, to be strategic in our work.

## Business analysis

Whether I am working in a government archive in Australia or a global corporation in San Francisco, one of the most important things I need to do is understand the people, the business and the records and data that is in my ‘jurisdiction’, and understand from a big picture perspective where I need to focus my efforts.

For example, as a government archivist, my job was influencing and providing advice on recordkeeping for the whole State of New South Wales in Australia. So it was important that my colleagues and I understood the functions of government, and where and how they were performed. This information, along with details of responsible agencies, people and systems, allowed us to intervene and assist as needed, to ensure adequate recordkeeping for business, accountability and memory purposes. I will talk more about that work in a minute.

In my current job, with a global fashion company, I spend much of my time analysing business functions, activities, systems and data, for the same reasons. However I am focusing mainly at present on the integration of privacy protections into these systems and processes, to ensure the company complies with regulations like Europe’s GDPR and a new law on privacy that is coming for California. Different contexts, but the methods are the same. I am still analysing business, risks and systems to decide on appropriate strategies for making, collecting and using records—it’s just they are records in data form, not ‘things’.

## Metadata design and implementation

I’d like to talk a little more about metadata—because it is so important! Metadata is fuel for the recordkeeping engine. Both librarians and archivists work with metadata, but, as I referenced earlier, in different ways.

Recordkeeping professionals—including archivists—work with metadata to ensure that records remain reliable and usable over time. We use metadata, for example, to ensure that a record:

- can be proven to be what it purports to be, because you can see metadata documenting its creation;
- can be linked to other records of the business it supports; and
- can be retrieved from systems of often millions of similar records.

We manage and add to metadata to maintain a record’s changing context, and to make it as usable (within the law and other requirements) as possible. Key standards for our work with metadata include:

- ISO 23081-1:2017 Information and documentation—Records management processes—Metadata for records—Part 1: Principles
- ISO 23081-2:2009 Information and documentation—Managing metadata for records—Part 2: Conceptual and implementation issues;
- (In draft) Records in Contexts (RiC): a standard for archival description developed by the ICA Experts Group on Archival Description

For archives, there are various models for the information we keep about our

collections. That last standard, *Records in Context* is an interesting development for Australians particularly in that it reflects an international change of direction with regard to archival description.

For example, archival traditions in North America and Canada have an approach that is based on the concept of the 'record group'. The record group is a way of describing any and all records retained as archives from a distinct entity, such as a government agency or a family. Under this model, an archive continues to add records to a record group up to the dissolution of the entity.

The National Archives and Records Administration (NARA) arranges its holdings under this model. So, as they state on their website, the records are described in such a way that they are: "attributed to the agency that created or maintained them and arranged thereunder as they were filed when in active use."<sup>2</sup>) This statement embodies the long established principle of provenance, or 'respect des fonds' whereby you keep records from the same creator together, instead of trying, for example, to divide them up according to subject matter. The last part about 'arranged as they were filed' is a reflection of the universal archival principle of 'original order', which was easy to implement in the paper world but not so clear cut for digital systems. If we have time I can return to that point in the Q&A part of the session.

So, NARA's system has the records creator as the central focus, usually a major government entity, such as a bureau or an independent agency. For example, National Archives Record Group 29 is Records of the Bureau of the Census. Most record groups include records of any predecessors of the organization named in the title of the record group.

A few record groups combine the records of several small or short-lived agencies having an administrative or functional relationship with each other. An example of this type of record group is Record Group 76, Records of Boundary and Claims Commissions and Arbitrations.

By contrast, since the 1970s, Australia has taken a different approach to the description of archives. One that still respects the principle of provenance, but enriches it. At the then Commonwealth Archives Office (now the National Archives of Australia), archivist Peter Scott proposed what would become commonly known as the 'series system'.

One of the characteristics of the Australian federal government at the time was a greatly increased rate of administrative change. Every time a new government was sworn in, they would create new departments, merge departments, rename and repurpose them. A descriptive system for archives that was reliant on the stability of records-creating administrative entities was simply not going to work. What was required was a system with flexibility, and resilience to change. This was what Scott devised.

The series system is, at its most basic, a three entity model. Think of it like a simple relational database.

The three entities are business (the functions, activities and processes that produce records), people and business units that do the business, and the records themselves.

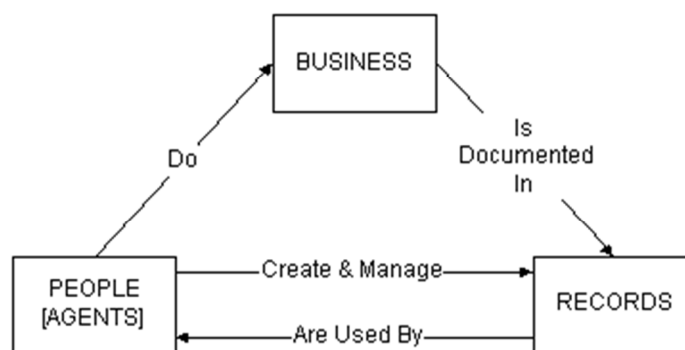
In the late 1990s, a research project at Monash University in Australia took the basic concepts of the series system and worked on a metadata model for all recordkeeping—regardless of the records’ context—being created in a business tomorrow, or being held by an archive for many years. This was called the ‘SPIRT’ project<sup>3</sup>.

Since then it has provided the basic design for all archives management systems in Australia—containing information about the records we hold.

These diagrams are from Monash University’s *SPIRT Recordkeeping Metadata Project—Conceptual and Relationship Models*, published in 1998, and they do a great job of explaining the basic building blocks of the Australian approach to metadata for records.

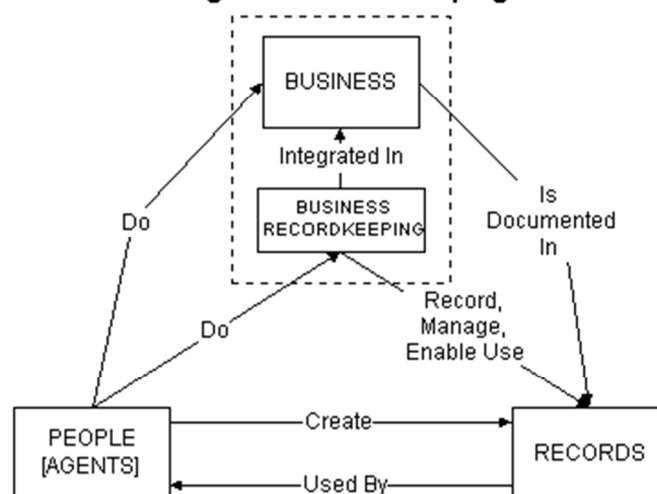
Figure 1 shows us the three main entities: Business, Agents and Records.

**Figure 1: The Business**



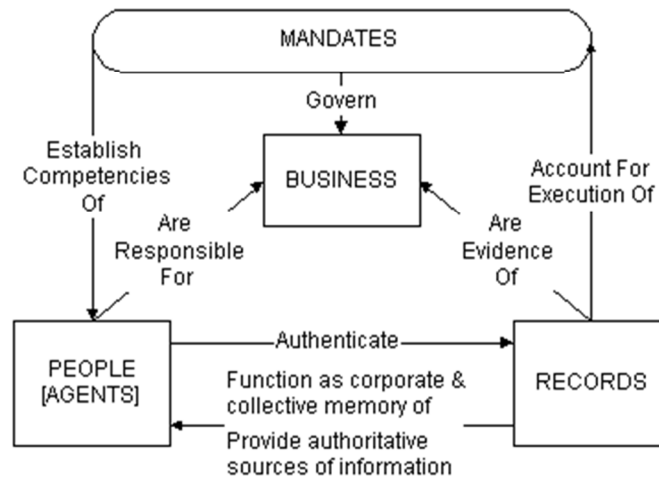
We then bring the systems for recordkeeping into the picture—here shown in Figure 2 as ‘business recordkeeping’.

**Figure 2: Recordkeeping**



And then we add the rules that drive and govern recordkeeping—mandates (laws, requirements for records).

**Figure 3: The Business Context**



So the work of description when you use the series system is about documenting these entities:

- Mandates (laws, other requirements affecting recordkeeping)
- Agents (People and business units)
- Business (Functions, activities and processes), and
- Records (from whole systems to individual units).

So you can see that the names of the departments that produced records at a certain point in time are recorded when we look at Agents, but the model does not rely on these in the same way that record group approaches do. In fact, as Australian practitioner and theorist Chris Hurley has observed<sup>4)</sup>, Scott’s work made the object of description the recordkeeping system as much (more even) than the records themselves. The Archives became a registry of recordkeeping systems, not merely a repository of archival ‘stuff’.

The beauty of the series system is the way it acknowledges change. Functions remain stable over time while the government departments that execute them change. Systems come and go, but the need for records of business persists. This model allows us to observe and document all this context, including rich relationships, in a way that is much easier to maintain than with hierarchical record group models. It is also very well suited to modern databases, to which new entities can be added as needed.

As an example, let’s look at just one entity: ‘Agents’. A government archive will probably register, or describe Agents from the highest level: the whole of the government, to government portfolios, such as ‘Environment and climate change’, down to individual agencies, like the ‘Office for clean air’. They will also register what we call ‘succeeding’ and ‘preceding’ entities, such as the department that came before the ‘Office for clean air’, which was known as the ‘Environmental



Protection Agency’, but was split into two in 2010, the ‘Office for clean air’ and the ‘Office for clean water’.

Undoubtedly, there will be further changes in the future. All of these entities produce records, which need to be linked to the appropriate office, so we can make sense of them now and in the future.

The functions are usually registered at a high level: ‘Environmental management’, for example, with subsets of Activities, like ‘Toxic waste monitoring’, or ‘Waterways management’. These, happily, remain much more stable than the Agent and Record entities. All of these must be related, both logically and in the systems we use, so that it is possible for an archivist or researcher to make the connections and find what they need.

A good example of how it all works is available from my old workplace, State Archives and Records NSW. If you like, you can have a look at the ‘advanced’ search options to see examples of all of these entities and how they are related to each other. Go to: <https://www.records.nsw.gov.au/archives/advanced-search>

Again, it is important to note that this approach is not limited to archival settings—we use it to manage recordkeeping in any context—including in current business settings. Much of the work I do at Gap Inc is about analysing and documenting agents (people, business units), mandates (laws affecting privacy and records retention) and records (collections of data made and kept by modern business systems that document business like recording customer preferences, or tracking online shipments).

So what does this all have to do with the *Records in Context* standard for archival description?

Well, it has adopted much of this model. The authors of the conceptual model for RiC have explained this as a move from multi-level to multi-dimensional description, in which the entities of records, agents and business are described at as many levels as is required. They are exploring the use of graph technologies to represent the model and eventually, for implementing archival description in practice.

I hope that you can see how the way we help people find records in archives is different from the way we help people to find books in libraries?

In archives we think about records creators and what they were doing when they produced the records, instead of thinking about the subject of the record.

I and other Australian colleagues, especially people like Barbara Reed, who developed Peter Scott’s work in the International Standards arena in the form of the metadata standards for recordkeeping that I mentioned are pleased to see this move towards harmonisation of descriptive approaches at the international level. Exciting possibilities exist for meta-description online and across national and sector-based boundaries, if the work of the ICA’s expert group responsible for RiC is widely adopted.

## Planning for change

The last aspect of our work that I would like to talk about is planning for change. Both business analysis and metadata design are vital to this element of our practice. As Peter Scott knew when creating the series system, when you are in archives and records, you are more attuned than many people to the inevitability of change. And one of the most important aspects of planning for change for recordkeeping professionals including archivists today is preparing for and managing systems migrations.

It is for these reasons that much of my career has been about working with IT and other partners to embed archival requirements into the plans for managing such systems, and about managing migrations of the systems as technology changes so the records they contain remain reliable and usable. This requires a knowledge of digital preservation tools and techniques, including metadata analysis, format identification and more. I find this to be one of the most interesting parts of recordkeeping and archival work, and one where we are seeing a lot of collaboration between archivists and librarians, on projects like Europe's Digital Preservation Coalition<sup>5)</sup> or more local groups like Australia's 'Australasia Preserves'<sup>6)</sup>.

When I was working for the New South Wales State archives, I led the team that built their first facility for accepting, preserving and making available the born digital records of government that were required to be retained permanently as archives.

In that work, we embraced the notion of the archive as a registry of recordkeeping systems as imagined by Peter Scott. Essentially, we considered the implications of this approach to keeping digital archives by reconceptualising the archival transfer process as consisting of system migration projects. Essentially this was about setting the right expectations with the transferring agencies. All organisations routinely migrate systems as a result of business or technological change. We argued that the challenge of moving digital recordkeeping systems to a digital archives is no different. Suggesting to agencies that the task is simply one of transferring digital objects makes it easy for them (their role is to just ship you a server or portable drive) but it can only work if you are willing to apply crude, universal solutions to the challenges of digital preservation and systems integration (i.e. how to intellectually merge an agency recordkeeping system into the digital archives). Framing the process as a systems migration makes it clear to agencies that they will need to resource these projects and be involved in their planning and execution. It also provides a useful trigger for transfer: migrations to digital archives can fit within broader projects to migrate legacy systems forward.

In the systems migration projects that we ran—and that the team continues to run—agencies collaborated with State Records to assess the particular needs of agency recordkeeping systems, agreed on customised preservation and access plans, and decided who was best placed to execute the elements of those plans based on levels of resourcing and expertise. This model incorporated preservation planning

within the process, rather than sitting it outside as a function solely for the archive and requiring general rules such as ‘convert all MS Office documents to PDF’—which may not have been necessary or appropriate. This is advantageous because it allows the characteristics of particular sets of records to be weighed in preservation decisions, it ensures agency buy-in for any format conversions that are necessary (so that agencies will remain confident in relying on those records post-conversion), and it is flexible and can cope with any type of record format or custom business system. Rather than being its focus, transfer (the formal transfer of custody) becomes just one element in a project and is no longer even a necessary element. While we were prioritising system migration projects that do involve the transfer of custody, the approach can be applied to non-custody migrations too. This means that, if an agency wishes to maintain control of its digital archives, the archival authority can still be involved in planning for the preservation and accessibility of those records.

A system migration approach treats recordkeeping systems holistically, maintaining their interrelationships and complexities and not treating them as just aggregations of files with a limited set of allowed metadata. It positions the digital archives program as a useful service that can support government agencies in doing digital recordkeeping. Migrations of digital records are happening more and more frequently in government and in the private sector, triggered by administrative changes and the shifting of functions between business units and even governmental jurisdictions. These are high risk activities that threaten the integrity and accessibility of digital records. A system migration approach for the digital archives means that the tools, practices and methodologies we develop can assist in migration work of all sorts, even when no archival records are involved.

## Conclusion

I hope that in this lecture I have given you an insight into what it means to work in archives and recordkeeping—at least what it means to me, as an Australian professional, albeit one who is now based in the US. I appreciate that there are different traditions in different parts of the world, but I firmly believe that as we move further into the digital age, it is archivists and other recordkeepers taking a proactive, strategic approach who will be best positioned to ensure the creation and availability of our documentary heritage into the future.

I would love to hear your questions, and perhaps in turn you could teach me a bit about archives and libraries in Japan. Thank you!

## References

- 1) Australian Society of Archivists, Position Paper: Archives and Libraries in the Digital Age, December 2017. Available at: <https://www.archivists.org.au/documents/item/1198>
- 2) National Archives and Records Administration, Record Group Concept <https://www.archives.gov/research/guide-fed-records/index-numeric/concept.html>
- 3) SPIRT Recordkeeping Metadata Project—Conceptual and Relationship Models <https://recordscontinuum.info/wiki/SPIRTRecordkeepingMetadataProject/ConceptualandRelationshipModels>
- 4) Hurley, Chris “What, if anything, is the Australian (“Series”) System?” <https://www.descriptionguy.com/images/WEBSITE/What-if-anything-is-the-series-system.pdf>
- 5) Digital Preservation Coalition <https://www.dpconline.org/>
- 6) Australasia Preserves <https://blogs.unimelb.edu.au/digital-preservation-project/2018/03/06/australasia-preserves-establishing-a-digital-preservation-community-of-practice/>

## Bibliography

Cassie Findlay <https://cassiefindlay.com/>

ISO 15489-1:2016 Information and documentation – Records management – Part 1: Concepts and principles

ISO 23081-1:2017 Information and documentation – Records management processes – Metadata for records – Part 1: Principles

ISO 23081-2:2009 Information and documentation – Managing metadata for records – Part 2: Conceptual and implementation issues

(In draft) Records in Contexts (RiC): a standard for archival description developed by the ICA Experts Group on Archival Description. Available at: <https://www.ica.org/en/records-in-contexts-ric-a-standard-for-archival-description-presentation-congress-2016>

Recordkeeping Roundtable <https://rkroundtable.org/>