Indigenous knowledge systems and linked data

ABSTRACT

This paper explores the development of a new framework to combine Indigenous knowledge systems and linked data to enable greater accessibility and culturally appropriate use of collection items within the GLAM sector.

The authors will discuss how incorporating Indigenous knowledge systems and understanding the complexities of linked data is important when classifying, cataloguing, and preserving knowledge for use by various audiences, from Indigenous community members to the general public.

Indigenous peoples have a wealth of intergenerational cultural knowledge passed verbally through the generations. This knowledge manifests in both intangible forms and material culture, including traditional knowledge and cultural expressions, oral traditions, dance, language, medicine, as well as tools and artworks. As a living culture, Indigenous peoples have knowledge systems that provide a greater understanding of their culture.

It is important that Indigenous communities maintain control over their cultural knowledge in order to preserve and share their knowledge in a culturally appropriate way. The GLAM sector would benefit through an arrangement with Indigenous custodians to incorporate Indigenous knowledge systems when determining linked data. Developing a better understanding of Indigenous knowledge and wider understanding of these cultures improves accessibility of connected resources to a wide range of audiences, in a culturally appropriate manner.

In addition to providing an overview of Indigenous knowledge systems, this paper also explores the description, retrieval and access to bibliographic and authority data using linked data principles in a GLAM environment. An introduction to linked data and the associated web standards, leads into discussion of the emerging BIBFRAME model. Linked data, with its basis in the semantic web, has the potential to deliver significant advances in discovering and sharing authoritative information to wider communities.

Possible benefits of using linked data for publishing descriptions and resources are explored, such as the ability to retain contextual relationships between items and collections. In addition, Linked data technologies can be used to expose the value added information about resources and their creators, such as bibliographic and authority data, to web search engines.

BIBFRAME is different to USMARC as it will provide a new way for collecting institutions and its users to annotate data by tagging and adding their own content. Cultural protocols and practices, such as the ATSILIRN (Aboriginal and Torres Strait Islander Library, Information and Resource Network) protocols and the

AIATSIS Guidelines for Ethical Research in Australian Indigenous Studies, are mentioned as an informed way to work with material with Aboriginal and Torres Strait Islander content.

PAPER

The Australian Institute of Aboriginal and Torres Strait Islander Studies (AIATSIS) is located in Canberra, on the shores of Lake Burley Griffin. For more than 50 years, AIATSIS has promoted understanding of Aboriginal and Torres Strait Islander cultures and heritage through supporting research, collecting materials and caring for and making accessible its dynamic and diverse collection. The AIATSIS Act was amended in 2016 to refocus the Institute's core activities on collecting, preserving and providing access to our national collection of Aboriginal and Torres Strait Islander culture and heritage.

A large collection of published and unpublished and unique material is held in our stacks and vaults consisting of over a million items. The collections include print and audiovisual material, personal and organisational records and reflect the diversity of Aboriginal and Torres Strait Islander Australia

AIATSIS has developed guidelines when conducting research with and about Torres Strait Islander peoples. Guidelines for Ethical Research in Australian Indigenous Studies[1] cover rights, respect and recognition, consultation and collaboration and appropriate use of and access to research. Ethical research respects Indigenous knowledge.

The key strategic priority for AIATSIS is to "ensure our collections are safe, accessible, valued and growing". [2] A major goal is to improve online discoverability and access to the collection. To do this we need to improve our systems as well as our management of the description of the collection. Crucially, we need to better understand and manage the rights – under law and under Indigenous Cultural Intellectual Property. One way to achieve the goal of discoverability and cultural richness of catalogues is to recognise Indigenous knowledge systems and include culturally appropriate descriptors.

This paper will explore:

- A definition of Indigenous knowledge systems
- The possible benefits of combining Indigenous and Eurocentric knowledge systems
- The imperatives to following cultural protocols
- Incorporating culturally respectful terms with linked data technologies

What are Indigenous knowledge systems?

The United Nations Educational Scientific and Cultural Organisation (UNESCO) describes Indigenous knowledge systems as the following:

Indigenous knowledge is the local knowledge that is unique to a culture or society. This knowledge is passed from generation to generation, usually by word of mouth and cultural rituals.[3]

Indigenous peoples have a wealth of intergenerational cultural knowledge that manifests in both intangible forms and material culture. This includes traditional knowledge and cultural expressions, oral traditions, dance, language, medicine and resources such as tools and artworks.

Indigenous peoples have knowledge systems that provide a greater understanding of their dynamic and living culture. It is well past time that Indigenous communities regain control over their cultural knowledge in order to preserve and share their knowledge in a way that they choose.

“However, older scientific methods of collecting and categorising Indigenous cultural materials has led to the items being taken out of context and redefined using Eurocentric methods of classification. This is problematic because it creates barriers to access for Indigenous peoples and perpetuates inaccurate or inappropriate representations for wider communities”.[4]

The International Federation of Library Associations and Institutions (IFLA) has released a set of guidelines titled ‘Statement on Indigenous Traditional Knowledge’ which encourages libraries and the information industry to assist with the “preservation and access to Indigenous knowledge.”[5]

As “Indigenous societies have a great many descriptive categories of the natural world, and their traditional knowledge and frameworks of thought cannot be separated from land, work, life-practices, and the relations between the individual or community and the natural world”[6], it is important to note that cultural meaning and understanding will be lost when interpreted in a different context, such as the Eurocentric classification systems in libraries.

Combining Eurocentric methods of classification with Indigenous knowledge systems

Are Eurocentric classification systems the best way to represent Indigenous knowledge systems? No, they are not structured in the same way.

In order to address these challenges, information professionals should acknowledge that “Indigenous knowledge systems are not exceptions or alternatives, and that Eurocentric methods of understanding and ordering the world are simply among several that have evolved in human societies.”[7]

Although collections within libraries and museums are a Western idea, by “navigating these collections and engaging with colonial records of the past, Indigenous people have been able to focus on the importance of connecting with the knowledge

contained within the records to revive, learn or relearn specific element of their culture."[8]

In addition to having different descriptive categories to the Eurocentric classifications used in collections, Indigenous knowledge also includes differing levels of cultural knowledge. These levels of knowledge and the accompanying access conditions may vary depending on the nature of the item – is it secret/sacred, culturally sensitive, or gender specific? Can that knowledge be shared with others not belonging to the Indigenous community? AIATSIS have established cultural protocols in order for information professionals to appropriately manage and provide access to cultural materials.

Library and museum practices do not naturally work well with Indigenous knowledge systems. Lawyer and leader in discussions of Indigenous Cultural Intellectual Property, Murri and Torres Strait Islander woman Terri Janke suggests that "the recognition of traditional knowledge rights will require coordinated systems that allow knowledge holders, intellectual property owners, creators and users to find a balance between sharing knowledge and cultural rights management."[9]

How does AIATSIS follow cultural protocols to share content?

The Aboriginal and Torres Strait Islander Library, Information and Research Network (ATSILIRN) Protocols is a guide for libraries and archives on good practice for those working with materials with Australian Indigenous content. The protocols acknowledge the moral right of Aboriginal and Torres Strait Islander peoples in being the owners of their knowledge. AIATSIS shows it's endorsement by aligning daily work practices with the protocols.[10]

It is important to share material with Aboriginal and Torres Strait Islander content in a culturally appropriate way. Rights to material covering aspects of Australian Indigenous studies can be complex. Our legislation requires us to consider appropriate access and as a consequence we have created an AIATSIS Access and Use Policy. This policy states "AIATSIS is committed to making its collections as accessible as possible, whilst respecting relevant laws and cultural protocols"[11]. This policy explains how AIATSIS provides access to and use of material in its collections on a daily basis within a best practice framework of Australian Law, the AIATSIS Act and Aboriginal and Torres Strait Islander cultural sensitivities.

Some materials held in libraries and archives may contain secret or sacred material. A sensitivity notice displays upon entry to the AIATSIS catalogue. In addition, each Aboriginal and Torres Strait Islander Index (ABI) record (an index to Australian Indigenous people mentioned in or as subjects in the published print collections at AIATSIS), contains a concise version of the sensitivity message.

In addition to the ATSILIRN Protocols, AIATSIS includes culturally appropriate terms within Pathways, the AIATSIS thesauri.[12]

The thesauri is a set of descriptors used for subjects relating to Aboriginal and Torres Strait Islander studies, placenames, language groups and people and reflects the content of the AIATSIS collections. The functions of the thesauri are to make searching for collection items in the AIATSIS catalogue easily accessible for Aboriginal and Torres Strait Islander community members, clients and researchers.

The thesauri are also used as a tool which AIATSIS cataloguers use to find appropriate terms to include as access points in catalogue records. The Pathways thesaurus has been recognised by the Library of Congress and is used in bibliographic records worldwide.

The Pathways thesaurus has been constructed following a Western model with a system of narrower and broader terms. The technical capabilities of linked data may allow us to show more appropriate relationships than the structures dictated by classical classification systems.

What is Linked data?

Linked data is about relationships. W3C says "The term Linked data refers to a set of best practices for publishing and connecting structured data on the Web."[13] The semantic web provides a standard way of expressing relationships. The Resource Description Framework (RDF) is a concept for data interchange on the web. These relationships are expressed as triplets. An example of this model is: ***Sally Morgan*** *Wrote a book called My place*

How does the web show these relationships?

Try Googling a noted anthropologist such as Professor William Stanner. A list of results is retrieved, and on the right hand side of the screen, a *knowledge graph* is displayed. Google describes a *Knowledge Graph* as "a term for how relationships are built between different people, places and things and reports facts about these entities."[14] The graph provides an information box as a combination of words and images. Biographical details are displayed along with an image of Professor Stanner and links to books he has written.

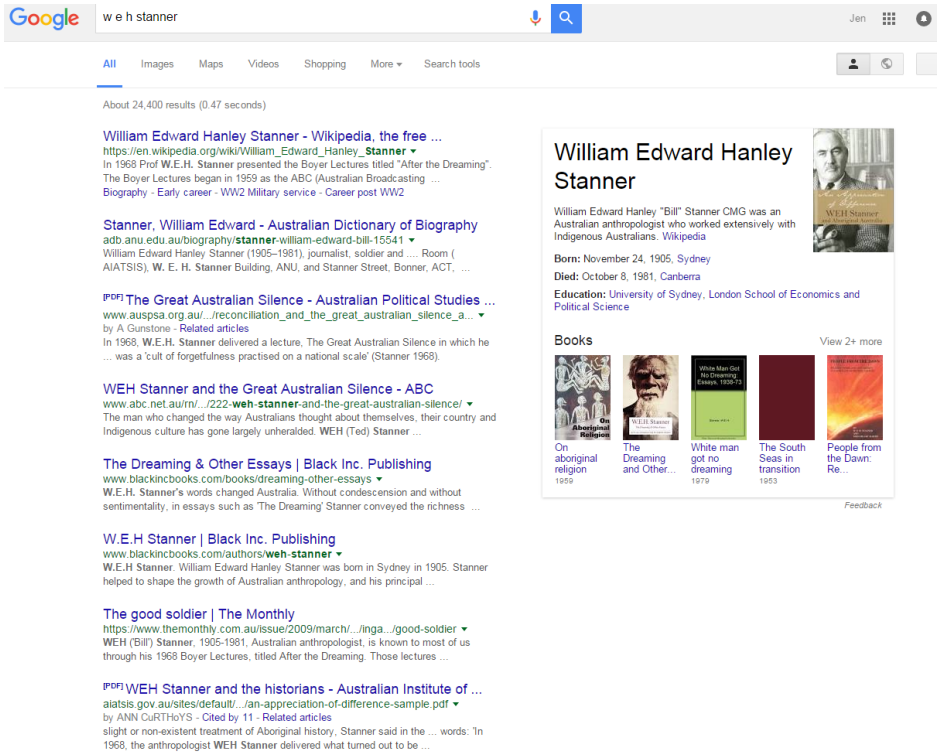


Figure 1 The Google knowledge graph is an example of linked data. https://www.google.com.au/#q=william+stanner

How is linked data of benefit to libraries and archives?

Libraries and archives need to have a presence where the users are. Many users start their literature searches in Google. What if a search on Google brought up an item held in a library or archive? An information box could display bibliographical details about the item and where it is held.

Why can't we do that now? Catalogue records are in a format called MARC, or MAchine Readable Catalogue, a standard devised over 40 years ago. MARC can't be understood by the web and doesn't support complex linking relationships between entities.

A new model is in development to replace MARC. The Library of Congress', BIBFRAME (Bibliographic Framework Initiative) uses linked data principals and the concepts of FRBR, (Functional Requirements for Bibliographic Records) and provides a new way to present catalogue records, reflecting the conceptual structure of resource materials. This is:

a Work – an intellectual or artistic creation

an Expression – the realisation of a work

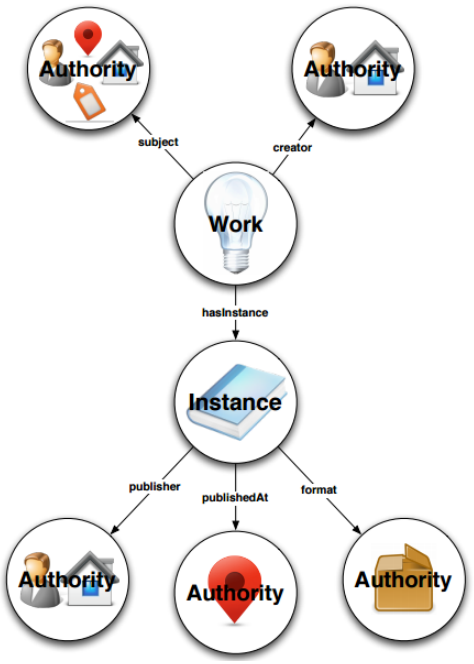
a Manifestation – the physical expression of a work

and an Item – a typical example of a manifestation[14]

Cataloguers started to hear these terms when implementing RDA (Resource Description and Access) – a content standard for describing materials. BIBFRAME will be compatible with RDA.

What is BIBFRAME?

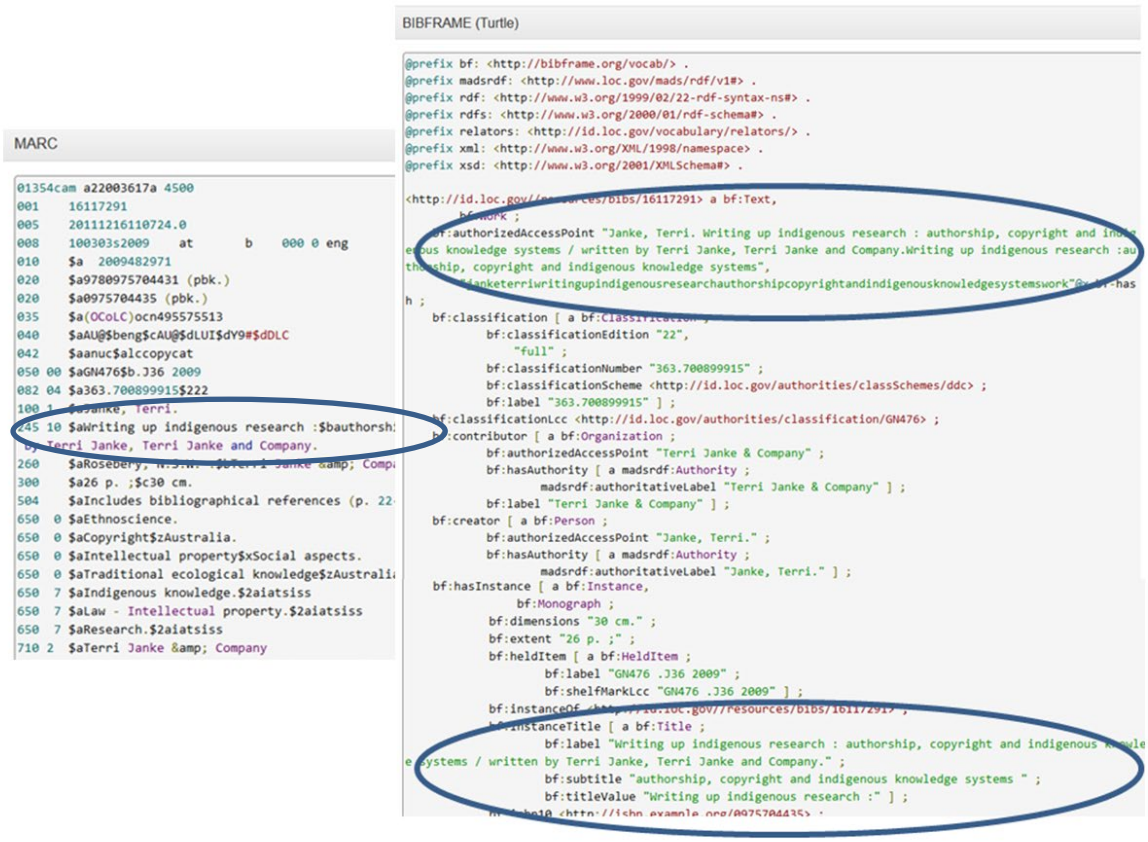
This diagram shows the conceptual model of BIBFRAME.



Conceptual model of BIBFRAME[15]

At the top are links to authorised subject headings and creator names. The 'work' is next, symbolised by a light bulb representing an idea. The 'instance' in this case is a book. The publisher, place of publication and format of the instance are linked to recognised authorised sources.

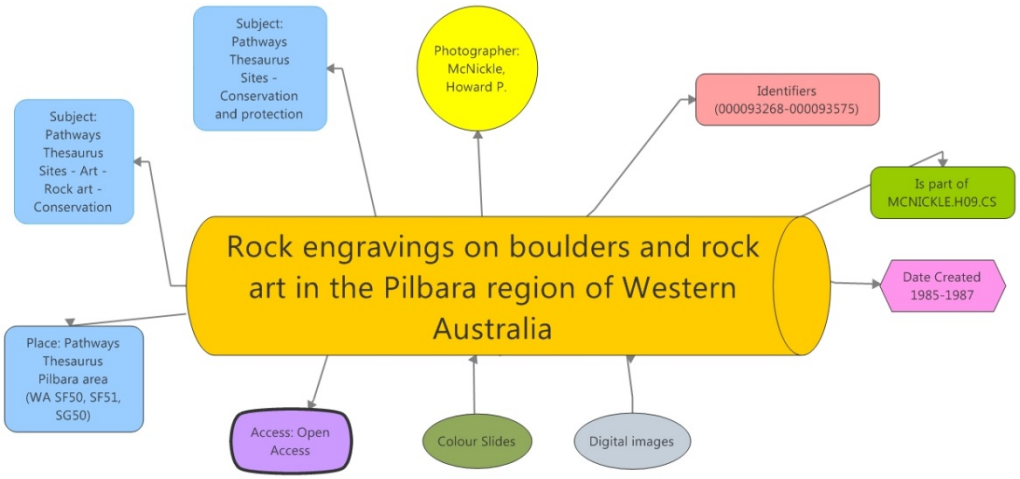
How does this look in a record? Below is a comparison of a book described using the MARC standard on the left and in BIBFRAME on the right.



Start with looking at the similarities in each record. Work and instance are circled in each record. Library professionals are quite used to seeing bibliographic information in the midst of MARC tags and subfield codes. What then are the differences? The data in a MARC record is comprised of strings of text; the data in a BIBFRAME resource points to entities or ‘things; ‘things not strings’ is a catchcry of linked data. This is moving from keywords that are used to describe things or entities to the entity itself. Different keywords can be used to refer to the same entity, whereas an entity is given a unique identifier. It will take time to adjust to seeing the bibliographic information amongst a different metadata schema with hypertext links throughout.

Libraries and archives have material arranged in collections. When an item is seen alone, it can lose the context of how and why it was created. An item may be described by a separate cataloguing record or be an individual entry in a finding aid. An overarching collection record gives these individual items context. Linked data is one way libraries and archives will be able to retain and enhance contextual relationships between items and collections.

We know that a variety of formats are held in collections and any descriptive tool needs to be able to be applied to all formats. Similarly, a photograph can be described using BIBFRAME. This example shows a BIBFRAME resource for photographs of rock art, whilst the creator is in this case a photographer, the content is the subject of Indigenous knowledge of which the community is the authority. Two instances of this photograph are described, colour slide and digital images.



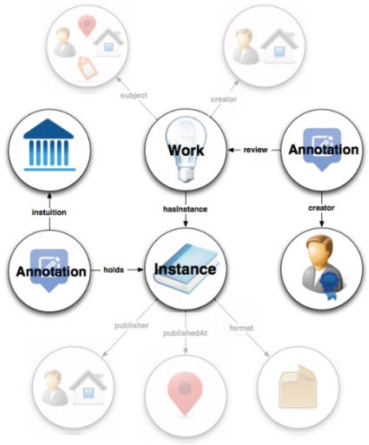
Example of a photograph from a collection held by AIATSIS

Linked data has the potential to deliver significant advances in discovering and sharing authoritative information to wider communities. Recently, AIATSIS delivered a file with 88,000 personal names from an authority file to the National Library of Australia to add to Trove, an aggregator of Australian bibliographic data. Through the Library, these authority records will feed through to VIAF (the Virtual International Authority File), hosted by OCLC, an American based cooperative for the library community to which Libraries across the world contribute. The National Library of Australia will include the authority records from AIATSIS in its next contribution. VIAF uses linked data technologies. A “super authority record” is created combining authority information for authors and people as subjects from around the globe.

Establishing appropriate personal name authority headings

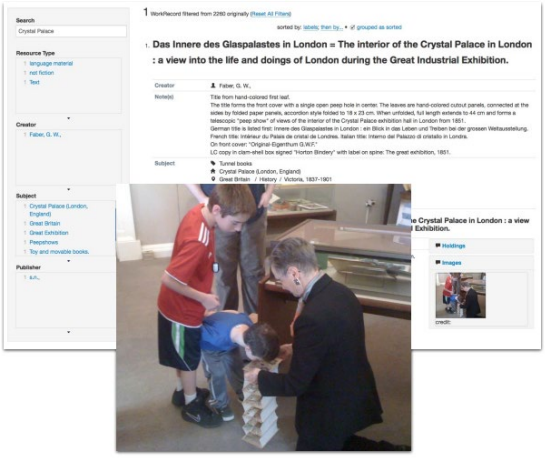
AIATSIS has developed a tutorial aimed at indexers and cataloguers working with materials in Australian Indigenous content. This tutorial gives an introduction to best practice for establishing headings for Aboriginal and Torres Strait Islander personal names. The tutorial includes self-paced modules and practice sessions for hands on experience The Talkin Names tutorial is a practical demonstration of the ATSILIRN protocols in action.

BIBFRAME can also include annotations.



Example of a BIBFRAME resource including annotation

Annotations are where libraries and archives can add value to the record – along with reviews, holdings, and local metadata. Annotations don't necessarily have to be textual. Below is an example of a user adding a photograph as an annotation to a BIBFRAME resource. An interface for addition of user generated content would need to be well designed to encourage participation and quality contribution.



Example of a user contributed annotation with BIBFRAME resource[16]

When will BIBFRAME be available?

BIBFRAME is a work in progress. The model and its components are still in discussion and development. However, there are already tools to transform MARC records into BIBFRAME resources. When BIBFRAME or any linked data model for libraries and archives is more mature, system vendors will need time to adjust to accommodate it. We can expect a mixed environment for some time.

Protocols and linked data

ATSILIRN Protocol number 5 Description and classification addresses the use of terms used in indexing, subject headings and classification. While these systems are designed to provide easy access to materials in libraries and archives, out dated or loaded terms can be a barrier to access. Protocol 5.5 encourages libraries and archives to "provide opportunities for Aboriginal and Torres Strait Islander people to describe and annotate material that relates to themselves and their communities."[10] The annotation part of BIBFRAME has the potential to enable libraries and archives to implement this guideline.

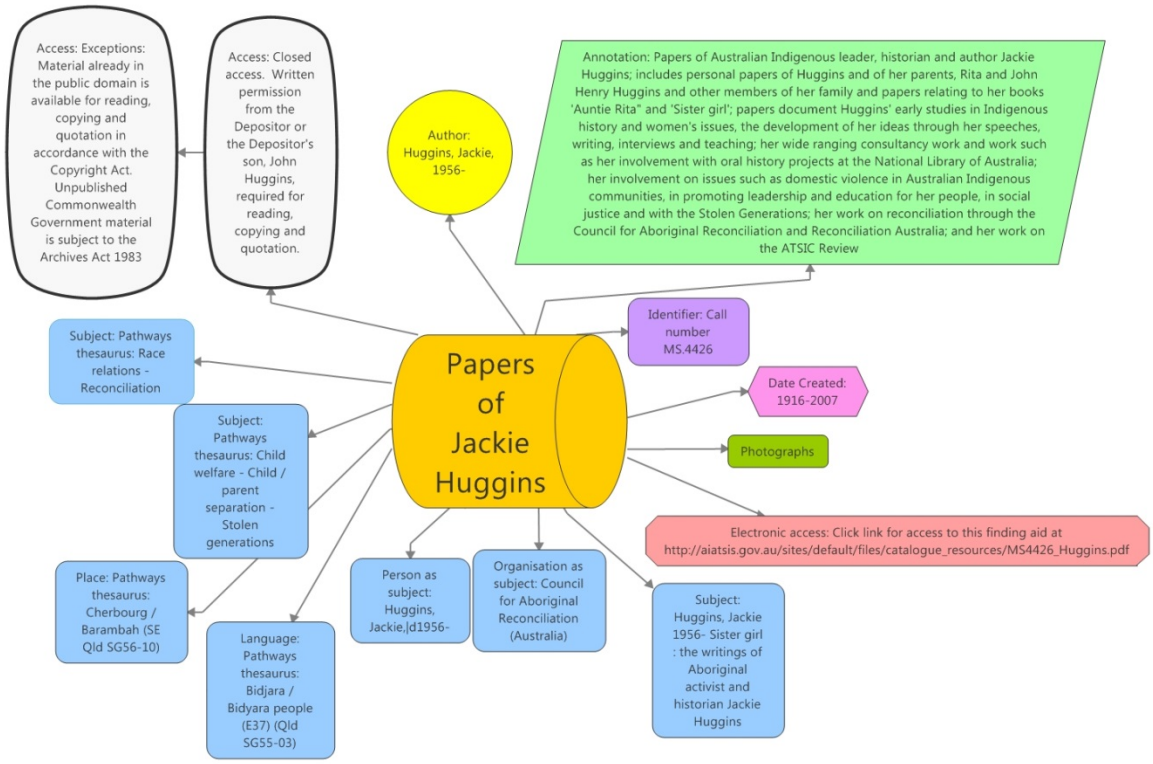


Figure: example of a conceptual model for a manuscript collection held by AIATSIS

AIATSIS is planning to update the publishing platform for the AIATSIS thesauri. An outcome of the review may be the ability to have the thesaurus published as linked data. This will give AIATSIS an opportunity to experiment with linked data projects, one of which could be to collaborate with Aboriginal and Torres Strait Islander peoples in discussing headings.

Strengths of Linked data include its ability to show relationships within data and that it uses 'things instead of strings'. Linked data technologies have been tentatively used by data scientists for more than a decade. Libraries and archives are new players in the linked data environment.

Our intention is to start a conversation about the challenges experienced by information professionals in providing access to Indigenous cultural materials. Also, to further develop methods of bridging the gap between Indigenous knowledge systems and Eurocentric systems of classification such as culturally respectful descriptors. Linked data offers one way of moving towards including Indigenous knowledge systems in recognition of different non-Western ways of maintaining and protecting knowledge. Technological advances may provide other ways into the future. Regardless the technology, code, platform, the fundamental issues are likely to persist and will require close collaboration with the owners of the knowledge systems that we seek to preserve and, as far as is acceptable to the owners, describe and share.

References

1. AIATSIS. "*Guidelines for Ethical in Research in Australian Indigenous Studies."* Canberra: AIATSIS, 2012. http://aiatsis.gov.au/research/ethical-research

2. "Corporate Plan - AIATSIS.gov.au." AIATSIS, n.d. Web. 25 Oct. 2016.

3. "Indigenous Knowledge Systems & Sustainability." UNESCO, n.d. http://www.unesco.org/education/tlsf/mods/theme_c/mod11.html

4. Ghaddar, Jamila and Nadia Caidi. "Indigenous Knowledge In A Post-Apology Era: Steps Toward Healing And Bridge Building". *Bulletin Of The American Society For Information Science And Technology*, vol 40, no. 5, 2014, pp. 41-45. *Wiley-Blackwell*, doi:10.1002/bult.2014.1720400511.

5. Whaanga, Hēmi et al. "He Matapihi Mā Mua, Mō Muri: The Ethics, Processes, And Procedures Associated With The Digitization Of Indigenous Knowledge—The Pei Jones Collection". *Cataloging & Classification Quarterly*, vol 53, no. 5-6, 2015, pp. 520-547. *Informa UK Limited*, doi:10.1080/01639374.2015.1009670.

6. Cherry, Alissa and Keshav Mukunda. "A Case Study In Indigenous Classification: Revisiting And Reviving The Brian Deer Scheme". *Cataloging & Classification Quarterly*, vol 53, no. 5-6, 2015, pp. 548-567. *Informa UK Limited*, doi:10.1080/01639374.2015.1008717.

7. ibid

8. Thorpe, Kirsten and Monica Galassi. "Rediscovering Indigenous Languages: The Role And Impact Of Libraries And Archives In Cultural Revitalisation". *Australian Academic & Research Libraries*, vol 45, no. 2, 2014, pp. 81-100. *Informa UK Limited*, doi:10.1080/00048623.2014.910858.

9. Janke, Terri. *Writing Up Indigenous Research*. Rosebery, N.S.W., Terri Janke & Company, 2009,.

10. "ATSILIRN - Aboriginal and Torres Strait Islander Library and Information Resource Network." *ATSILIRN - Aboriginal and Torres Strait Islander Library and Information Resource Network*. N.p., n.d. Web. 13 Aug. 2016. http://atsilirn.aiatsis.gov.au/protocols.php

11. "Access and Use Policy - Australian Institute of Aboriginal Studies" AIATSIS, n.d. Web. 25 Oct. 2016.

12. AIATSIS. "Pathways: The AIATSIS Thesauri Homepage." *Pathways: The AIATSIS Thesauri Homepage*. AIATSIS, n.d. Web. 13 Aug. 2016. http://www1.aiatsis.gov.au/

13. "Best Practices for Publishing Linked Data." *Best Practices for Publishing Linked Data*. N.p., n.d. Web. 13 Aug. 2016. http://www.w3.org/TR/ld-bp/

14. "Google Launches Knowledge Graph To Provide Answers, Not Just Links." *Search Engine Land*. N.p., 2012. Web. 13 Aug. 2016. http://searchengineland.com/google-launches-knowledge-graph-121585

15. "FRBR." *FRBR*. OCLC., n.d. Web. 13 Aug. 2016. http://www.oclc.org/research/activities/frbr.html

16. Library of Congress. "Bibliographic Framework as a Web of Data: Linked Data Model and Supporting Services." Library of Congress, 2012. Web. Developed under contract from the Library of Congress by a team from Zepheira composed of Eric Miller, Uche Ogbuji, Victoria Mueller, and Kathy MacDougall. http://www.loc.gov/bibframe/pdf/marcld-report-11-21-2012.pdf