Linked Data in Small Archives

A Case Study of the David Fanshawe World Music Archive

KIRSTY MORGAN

This dissertation was submitted in part fulfilment of requirements for the degree of MSc Information and Library Studies

Dissertation Supervisor: Dr Diane Pennington

18 AUGUST 2019

Word Count: 21,983

DECLARATION

This dissertation is submitted in part fulfilment of the requirements for the degree of MSc of the University of Strathclyde.

I declare that this dissertation embodies the results of my own work and that it has been composed by myself. Following normal academic conventions, I have made due acknowledgement to the work of others.

I declare that I have sought, and received, ethics approval via the Departmental Ethics Committee as appropriate to my research.

I give permission to the University of Strathclyde, Department of Computer and Information Sciences, to provide copies of the dissertation, at cost, to those who may in the future request a copy of the dissertation for private study or research.

I give permission to the University of Strathclyde, Department of Computer and Information Sciences, to place a copy of the dissertation in a publicly available archive.

(please tick) Yes [✔] No [ ]

I declare that the word count for this dissertation (excluding title page, declaration, abstract, acknowledgements, table of contents, list of illustrations, references and appendices) is **21,983**.

I confirm that I wish this to be assessed as a Type 1 (2) 3 4 5

Dissertation (please circle)

Signature:

Date: 18/08/2019

ABSTRACT

Libraries, archives and other cultural heritage institutions are trying to become more connected and there is concern that smaller collections will become increasingly overlooked in the future. This dissertation examines the case of the David Fanshawe World Music Archive – a small but significant collection of sound recordings, photographs, journals and other accompanying material – to explore the feasibility of the Archive implementing some kind of Linked Data system to make it more accessible online. Through a case study of the Archive and interviews with its staff and with four other libraries, the research develops a wider understanding of the current situation in music libraries and archives and their readiness for developing accessible online presences. Online catalogue creation and digitisation projects are discussed.

The dissertation explores Linked Data as a key element in the creation of the Semantic Web/Web 3.0. It highlights the processes and concepts involved in Linked Data but finds that libraries in general and small libraries and archives in particular, are a long way from being able to use Linked Data in their catalogues. Nevertheless, the steps which are being taken to making collections available online are explored. These are found to be potentially valuable in themselves and to be important stepping stones which, if set in place appropriately from the start, may be essential in allowing an easier transition to online accessibility and ultimately to engagement in the Semantic Web at a future point, if smaller institutions choose to go down that route.

The complementary roles of librarians and tech specialists are highlighted. While some librarians and archivists may be wary of the technical aura surrounding Linked Data, their expertise in areas such as vocabulary control, ontologies and cataloguing are essential if the potential of the Semantic Web is to be fulfilled.

The dissertation makes specific recommendations for the Fanshawe Archive relating to its current and future projects and draws wider conclusions for the library and cultural heritage sectors in general. Wider recommendations include collaboration amongst institutions and the community of users, and a role for organisations such as CILIP, IAML and the MLA in promoting a librarian-friendly Linked Data OPAC system and a World Music focused ontology.

KEY THEMES

Cataloguing -- Digitisation -- Linked Data -- Music Libraries and Archives -- Ontology -- Semantic Web

ACKNOWLEDGEMENTS

I would like to thank my dissertation supervisor, Dr Diane Pennington, and everyone who participated in the interviews. Particularly, thanks to the Fanshawe Archive for their role as the case study and the Fanshawe Manager for her hospitality.

TABLE OF CONTENTS

TABLE OF FIGURES

INTRODUCTION

Libraries, archives and other cultural heritage institutions are trying to become more connected. The British Library, for example, is examining the possibility of a server that connects public library catalogues to help patrons and staff find out what libraries outside their districts have in their collections (Connolly *et al*., 2019). There are concerns about libraries being information silos, with many valuable and interesting resources that few members of the public know about (Thorson & Pattulli, 2016). One way that some libraries are trying to be more available online is to use Linked Data (LD).

Some libraries, including the British Library, are already using LD to make their collections more online accessible, theoretically pushing search engine results about their collections higher. If already well-established libraries are using this tool to reach a wider audience, could this mean that smaller collections will be further overlooked? This dissertation examines the case of the David Fanshawe World Music Archive – a small but significant collection of sound recordings, photographs, journals and other accompanying material – to explore the feasibility of the Archive implementing some kind of LD system and make recommendations for this archive if it were to choose to go down that path. The dissertation considers lessons from the Fanshawe Archive's specific case, which may be more widely applicable. Interviews with other libraries broaden the research to develop a wider picture of the current situation in libraries and their readiness for developing accessible online presences.

RESEARCH QUESTIONS

Main Question: **How could the David Fanshawe World Music Archive use Linked Data to make the Archive more accessible online?**

The project's overall aim is to carry out a Case Study of the Fanshawe Archive, to determine whether LD could be implemented to make the Archive's collection more accessible online. The project hopes to provide answers to the following sub-questions to answer the main question above:

1. What are the Fanshawe Archive's aims for its collection, in terms of target audience, accessibility, wider appeal and future vision?

The dissertation begins with a description of the Fanshawe Archive and its current situation. Chapter 1 uses interviews that were conducted with the Fanshawe archivists alongside literature about David Fanshawe to provide a picture of the Archive, including a description of the target audience, the Archive's online presence, their aims for the collection, current cataloguing and digitisation projects and their future hopes for the Archive.

2. What are the best practices for creating LD, particularly in a music archive?

Chapter 2 reviews current literature into how the library community is, and could be, using LD. It describes recommended best practice for creating LD and explores LD in the contexts of libraries, small archives and music. It also briefly touches on the ethics of Western institutions sharing materials from non-Western cultures, since this is particularly relevant for the Fanshawe Archive, as a World Music collection.

3. What progress has the Fanshawe Archive made with its current projects to catalogue and digitise its analogue materials?

Chapters 1 and 4 discuss the Fanshawe Archive's current cataloguing and digitisation projects, drawing on interviews, a site visit and documentary analysis of the Archive.

4. Does the Fanshawe Archive have the resources (technology, staff, budget etc.) to realistically implement LD?

Chapter 4 examines the Fanshawe Archive's projects and resources, using interviews with the Archive and other aspects of the site visit, including documentary analysis, to appraise whether the Archive realistically has the resources to support an LD project. Interviews with librarians from outside the Fanshawe Archive inform the understanding of resource requirements for LD and similar projects.

5. How does the Fanshawe Archive operate now, and what could they do to implement an LD system?

Chapters 1 and 4, through analysis of interviews and Archive documents, describe what the Archive might need to add or change to implement an LD system. Interviews with libraries outside the Archive (an anonymous national library that has already started to use LD, and three smaller music libraries with similarities to the Fanshawe Archive) are used to explore what other institutions within the music heritage sector are doing, and see whether any inspiration can be drawn from their practices. Chapter 5 builds on the interviews, literature, and discussions in previous chapters to create an LD model using digitised files from the Archive, to demonstrate how the Archive might begin implementing an LD system, if or when they choose to do so.

6. How could the Fanshawe Archive's situation apply to a wider context?

Although the case study focuses on LD through the lens of the Fanshawe Archive, there is broader relevance for other libraries and archives, particularly within the music field. The interviews with staff from four other libraries place the Fanshawe Archive into a wider context and are discussed mainly in Chapters 4 and 5. The interviews inform discussion on previous projects to make libraries more accessible online, particularly within the Linked Datasphere. The Fanshawe Archive may

choose not to implement LD, but Chapter 5's LD creation process discussion has the potential to act as a starting point for other institutions. Some of the recommendations in Chapter 6 will also be transferable for other cultural heritage institutions.

Ultimately, the project hopes to answer the main question: **How could the David Fanshawe World Music Archive use Linked Data to make the Archive more accessible online?**

CHAPTER 1 – DESCRIBING THE FANSHAWE ARCHIVE

1.1.What is the Archive?

The Fanshawe Archive was created by composer and ethnomusicologist David Fanshawe. After he died in 2010 it was placed into the David Fanshawe World Music Archive Trust. It is now run by his wife (FM) and two other trustees (FT(L) and FT(D)), with one part-time member of staff (FPT). It is a very personal collection, maintained by people close to Fanshawe, so I feel it is exceptionally important to respect their wishes and vision for the Archive when discussing their situation.

1.2.A Brief History

Fanshawe trained as a sound engineer and despite not attending music college he had a profound interest in sound, including World Music (FM, 2019). Inspired by ethnomusicologists from previous generations, including Vaughan Williams, Percy Grainger and Hugh Tracy, Fanshawe spent eight years in the Middle East and Africa recording indigenous music (McCarthy, Date and Fanshawe, 2018). From the outset, he kept logical, detailed records of his recordings (FT(L), 2019), including where the recording took place, who was performing and why, and attributing credit to the performers wherever possible (FM, 2019). He was a skilful sound engineer and had the expertise to get quality recordings despite the remote conditions and limitations of analogue tapes. The recording quality makes the collection particularly significant.

Fig.1.1. Fanshawe in his Archive
©David Fanshawe

Before his travels Fanshawe was a composer, and after his Africa trip (1969-73), he returned to composition for four years, composing one of his most famous compositions, *African Sanctus*, which combined some of his African recordings with a Christian Mass. FT(D), who helped Fanshawe with editing recordings for *African Sanctus*, remembered working on the 'Call to Prayer' movement until 4am. In the morning, the twin boys surprised them from the upstairs landing with a performance of the 'Call to Prayer', which had kept them awake all night (FT(D), 2019)!

Eventually, overwhelmed by the pressures of composing, Fanshawe "spun the globe" to see where he could travel next and settled on the Pacific Islands (FM, 2019). He created 2,000 tapes-worth of recordings in a bigger expedition than his previous 500-tape African journey. Afterwards, he composed *Pacific Odyssey*, something of a sequel to *African Sanctus* in that it also combined indigenous recordings with traditionally European sounds.

Fanshawe's passion for World Music shines throughout the collection. The high-quality recordings, breadth of music, and accompanying photographs and journals make it an important piece of research. Fanshawe sadly died in 2010, but the Archive aims to preserve and share his collection. Perhaps Linked Data could help them promote this fantastic resource.

1.3.Archive Layout and Contents

Fig.1.2. The David Fanshawe World Music Archive

Fanshawe set up the Archive in Australia and then in Wiltshire. After his death the Archive was moved to the temperature- and humidity-controlled basement of FM's current home in Malvern – no easy task (FT(D), 2019). It looks fantastic: 3,000 tapes line the walls according to Fanshawe's original classification scheme. 1,000 boxes of coloured slide photographs are kept in labelled drawers and can be looked at through a slide viewer just outside the room (Fig.1.3.). Half a wall is dedicated to Fanshawe's travel journals, with reference books about the places Fanshawe visited. Alongside all this, there are boxes of ephemera, including letters, indigenous clothing such as a collection of Maasai necklaces, Fanshawe's original recording equipment and even a box for Fanshawe's distinctive caps! FM has also donated some native instruments that Fanshawe collected on his travels (FM, 2019).

While this dissertation focuses on the Archive from a digital online perspective, the physical archive is almost an artefact in itself. Although many World Music sound recordings exist, it's rare to find so much complementary material, particularly the high-quality photographs. Everyone involved with the Archive is keen for the collection – sound recordings, photos and journals – to physically remain together (FM, 2019; FT(D), 2019; FT(L), 2019; FPT, 2019). One of the Archive's long-term goals is therefore to find an academic institution wealthy enough to house the collection but small enough to "cherish" it for the significant collection it is (FM, 2019).

Fig.1.3. Fanshawe Archive's Slide Viewer

The Archive's cultural value is further enhanced by the fact that many of the recordings would be impossible to record today. In a BBC Documentary retracing his African Journey, Fanshawe discovers that many of the performers are no longer alive (Fanshawe *et al.,* 2006). Some died of natural causes, others were victims of dictatorial regimes; the Bwala Dancers from Uganda, for example, had been murdered by Idi Amin's army (FT(L), 2019). Recordings from other areas have a similar story. The spirit song 'U'Ula' from Melanesia was performed by four older ladies who have since died and the tradition would be "difficult to record today" (Fanshawe, 1998). There are also recordings and photographs of the Iraqi Marsh Arabs, who were severely persecuted by Saddam Hussein.

1.4.Staff

The four people involved in the Archive don't have official roles; rather, they help out wherever they can. However, they each provide complementary skills to the Archive.

FM has a great knowledge of the Archive, having been highly involved with it for a long time, both while Fanshawe was alive and later managing the Archive and working towards its current position. She is also familiar with the music and cultures that Fanshawe recorded, particularly the Pacific areas, since she was married to Fanshawe during his Pacific journey.

FT(D) is a former sound engineer with decades of experience and worked closely with Fanshawe. He has been the main actor in the tape digitisation and supplies his knowledge of sound and recording to the team.

FT(L) is a former librarian who has worked as a cataloguer. Passionate about the Information side of IT, maintaining professionalism in the library sector, and cataloguing properly and consistently, FT(L)'s expertise will be useful in the creation of a Thesaurus, selecting terms to use, supporting the catalogue creation and deciding "where to stop" (FT(L), 2019) with resource description.

FPT is the main support in maintaining the Archive day-to-day, alongside FM, and has created a template for how the catalogue interface might look. Very active in the catalogue creation, FPT has an interest in information science and researched LD in advance of our interview.

Each brings useful expertise that helps them collaborate to maintain the Archive and carry out necessary projects.

1.5.Archive's Current Reach

At the moment, the Archive isn't very heavily used but they are now looking to increase their reach. Currently, the Archive mainly caters to music students, particularly those interested in niche aspects of World Music, such as North African desert harps or South Polynesian gospel hymnary chanting; and to commercial users: for example, providing material for World Music compilation albums or filmmakers looking for authentic recordings (FM, 2019). FM and FPT deliver talks about World Music and David Fanshawe's collection to local schools and the music is used in exhibitions in England and abroad (FPT, 2019). They also offer guest speaking at universities and have produced a vision brochure about the Archive that they will send to universities and libraries across the UK (FT(L), 2019; FM, 2019).

The Archive currently has no real online presence. FM runs a Fanshawe Music Business, completely separate from the Archive and trust, that sells CDs in an online store, but so far, the Archive itself is not mentioned on the website or anywhere else online. There are plans to revamp the Fanshawe Music website with a branch for the Archive (FM, 2019), but this has been placed on a backburner with the trust more focused on creating a catalogue and digitising the collections (FT(L)), discussed below and in Chapter 4. In the longer timeframe, however, the Archive does plan to put some of the collection online, in accordance with their aim to make the Archive as wide-reaching as feasible (FPT, 2019). FT(L) (2019) stated, "I just think he'd love to think that his work was available to everybody, every age, all ages, all countries, all people." Making the collection available online would help the Archive to remain true to David Fanshawe's vision for his remarkable recordings.

1.6.Digitisation

Before the Fanshawe Archive can put its material online, they need to make the collection available in a digital format. They are currently in the midst of a big digitisation project to preserve the

Fig.1.4. Magnetic Tape Player

collection as well as transfer it to a more accessible format. They are scanning all Fanshawe's notes, from the journals and the original "Masterboxes". They are also digitising the photograph slides, which are at risk of decolourisation (FM, 2019), and analogue tapes, which are in a generally good condition but fragile and therefore not available for people to listen to (FPT, 2019). Some tapes will require extra work due to "Sticky Tape Syndrome", where tapes from the 1980s excrete a sticky substance that makes the tape stick and the digitisation process take much longer (FM, 2019). The Archive has a baking oven onsite to heat sticky tapes so they can be digitised, but the process is very slow, and the baked tapes must be digitised that day (FM, 2019).

FT(D) generally takes batches of tapes home to digitise, but during the site visit he set up a digitisation station and demonstrated how it worked. The sound is amplified by an amplifier and processed using the free software program Audition 3. Each 22-minute tape is played in real time and needs someone to be listening for any hiccups in the transfer process. FT(D) explained the analogue tapes naturally shed iron oxide, and if there is a build-up then the receiver stops registering the higher frequencies. Some tapes have also been edited and restuck with Sellotape, which can come loose and make the tapes skip as they're being transferred.

The Archive has already paid for some slides to be digitised, but they are planning to invest in a slide scanner (£500-£1,000) to do the scanning themselves. This has three advantages: for the number of slides it's cheaper to buy a scanner than to pay for external digitisation; it gives them complete control over the scanning process; and the slides wouldn't have to be transported, which lessens the risk of damage. The main drawback is the lack of staff to do the scanning, which isn't really automatable. According to FT(D) each scan takes 30-45 seconds, and then the slide needs to be manually switched to the next one. This makes it difficult to do any other work during the scanning, although perhaps scanning and digitising the tapes could be done at the same time, if the equipment were set up in the same room. FT(D)'s demonstration during the site visit indicated that although the digitiser had to be listening the whole time, the hands-on work was intermittent. A relatively quiet slide scanner, and a good set of noise-cancelling headphones

Fig.1.5. Digitisation Station

(around £300) would ensure that the digitiser could still hear the tapes well enough to quality-check them while scanning the slides and this would make the digitisation process quicker. The same setup could work when scanning the journals. However, since FT(D) said that the tape digitisation is a "specialist process", it might be a more complicated and hands-on task than it appeared to me as an outsider.

The Archive has arranged to send 600 tapes to be digitised by the National Heritage Lottery-funded and British Library-supported project, Save Our Sounds, which should complete the Africa collection and make a good start on the Pacific collection. FM (2019) told me that since Fanshawe took such care in his recordings, the Archive is not digitally enhancing the tape quality in any way. They have asked the Save Our Sounds project to refrain from digital enhancement also. This returns to the Archive's purpose to preserve Fanshawe's legacy and the personal nature of the collection, since Fanshawe's skill of recording the sounds to such high quality is part of what makes the collection special.

1.7.Catalogue Creation

Currently, the Fanshawe Archive's catalogue is a printed Word document listing the tapes by year recorded. Like most home catalogues, this might work for people already familiar with the collection but for external users it's not ideal. The Archive recognise the need for a more user-friendly catalogue and have plans to transfer it to an Excel file, which will be easier to search (FPT, 2019). FPT has created a template for how an ideal catalogue interface might look, arranging the metadata into four categories. Technical Data includes the recording date, recording length and notes about the recording equipment. Cultural Data is information about the place the sound was recorded and the language of the recording. Musical Data covers the reason for the recording (e.g. a wedding), the performer, the type of performance and performance description. Archival Data includes a transcription of Fanshawe's Masterbox notes, links to the journal and photo references, and the recording's physical location in the Archive. This list isn't exhaustive, but it offers a brief overview of the types of things in each category.

Interestingly, again in the spirit of preserving Fanshawe's archive as he intended, the Archive is clear that they want to include Fanshawe's original tape numbers and original information. They acknowledge that this might not be the most crucial information for the general public, but feel ethnomusicologists, for example, might be interested in it (FPT, 2019). FT(L) (2019) told me that it would be possible to include catalogue fields that would aid access as well as fields for the original tape numbers.

The catalogue is still in its early planning stages, but the Archive's thoughts on what to include in a catalogue will be useful for me when discussing how they could use LD.

1.8.Fanshawe Archive's Initial Thoughts on Linked Data

Right now, taking on an LD project is not a priority. The Archive are concerned that LD would be expensive, require expertise they don't have, and they are already time-stretched with their current projects. But they also would like to make their archive more accessible in the long run, and if LD could help them do that, then all four have indicated that it might be something to explore in the future. FT(D) (2019) mentioned that FPT might be particularly willing to get involved due to her interest in information science. This project therefore examines whether there is anything the Archive can do now to make using LD easier in future and how they might go about tackling LD in a few example cases if they were to ever implement such a project.

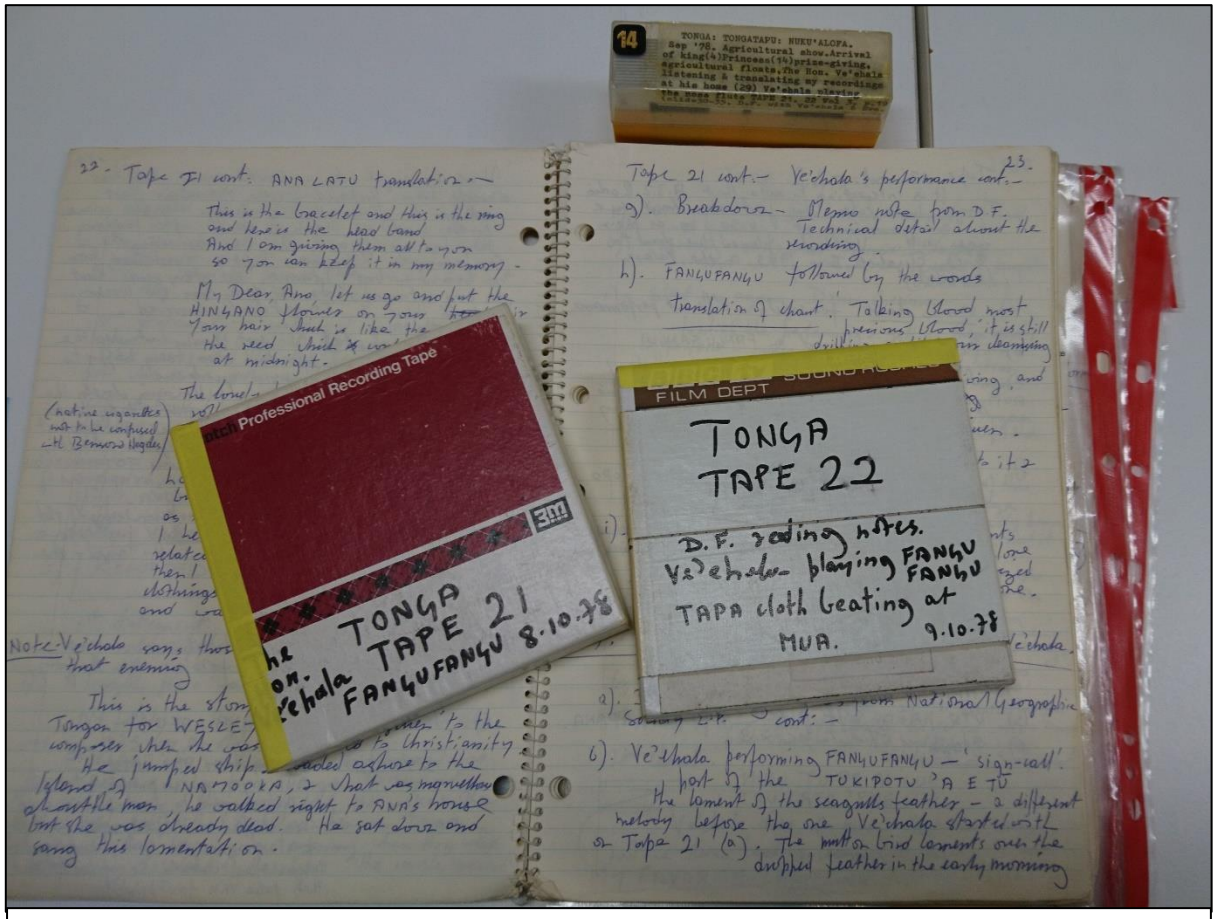

Fig.1.6. The Archive has two tapes, several journal entries and almost an entire box of slides devoted to the Fangufangu Nose Flute (A digitised photo of from one of the slides can be seen in Chapter 5).

CHAPTER 2 – LITERATURE REVIEW

2.1.What is Linked Data?

LD is a way of structuring online metadata so that computers can better understand the semantic relationships between items. LD tools and metadata structures help move towards the ultimate goal of creating the 'Semantic Web', also known as Web 3.0 (Pennington, 2016), where computers are able to interpret information on the web to automatically deduce relationships between resources and thereby enrich the users' browsing experience by helping them discover related sources they might not have otherwise found (Pennington and Cagnazzo, 2019). Cagnazzo (2019) clarifies that, if the Semantic Web refers to the end vision of an online network that computers can properly understand, then LD is how that end vision can be achieved.

To understand LD, let's take it back a step. A regular hyperlink between two documents online can tell a computer that the resources are connected in some way, but the computer cannot understand how the documents are related or the reason for the hyperlink (Sporny, 2012). To a computer, one hyperlink is identical to any other hyperlink. LD uses Resource Description Framework (RDF) to structure metadata about an online item so the computer can read it; importantly, RDF defines the relationship between the item and other resources online (Pennington, 2016). It does this by using RDF "triples", which are structured to correspond to the different parts of a sentence: Subject → Predicate → Object, for example, Franz Schubert → composed → Winterreise (Shotton, 2013). Once coded into XML following the RDF structure, the intermediate step (the predicate) provides the computer with the important information about the nature of the relationship in a machine-readable format (Pennington, 2016). So, if provided with the additional triples:

Beethoven → composed → Fidelio
Franz Schubert → lived in → Vienna,

the computer would be able to tell that the relationship between Schubert and Winterreise is the same as that of Beethoven and Fidelio, but different from the relationship between Schubert and Vienna, since the predicate in the latter triple is different. This allows a searcher to filter their search better, letting search engines retrieve more relevant results (Fay and Sauers, 2012). The triple, also called a statement, is directional, and should therefore be reciprocated (Pennington, 2016):

Winterreise → has composer → Franz Schubert

Additionally, RDF uses Unique Resource Identifiers (URIs) to describe items and relationships for the computer; these are generally the resource's http web address (McCallum, 2016) or namespace (Bizer *et al*., 2009). Assigning each item a unique identifier helps maintain consistency and prevent

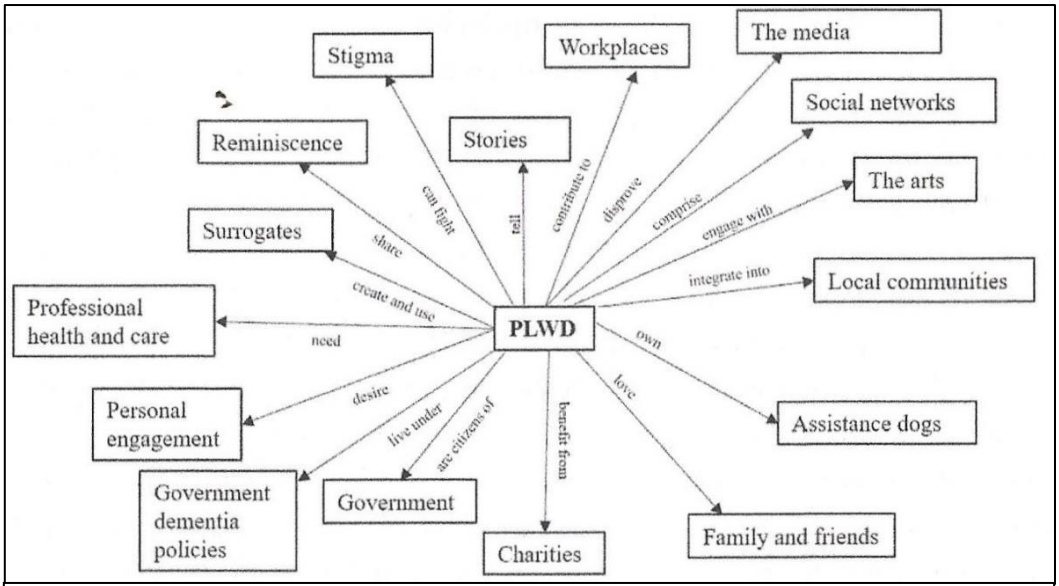

*Fig.2.1.*Concept Map of the start of a Dementia ontology. PLWD stands for "person living with dementia". (Pennington, 2019).

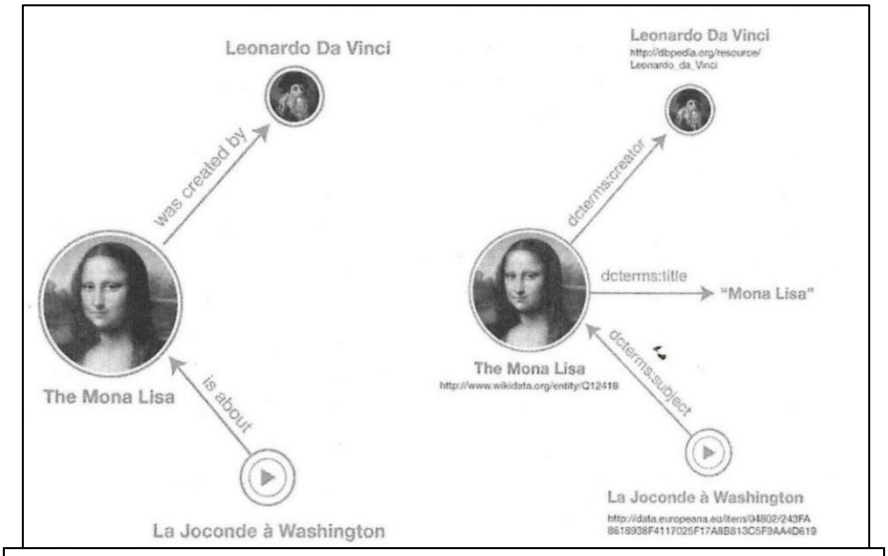

*Fig.2.2.*Transition from human-language to computer-language. (Syn, 2019).

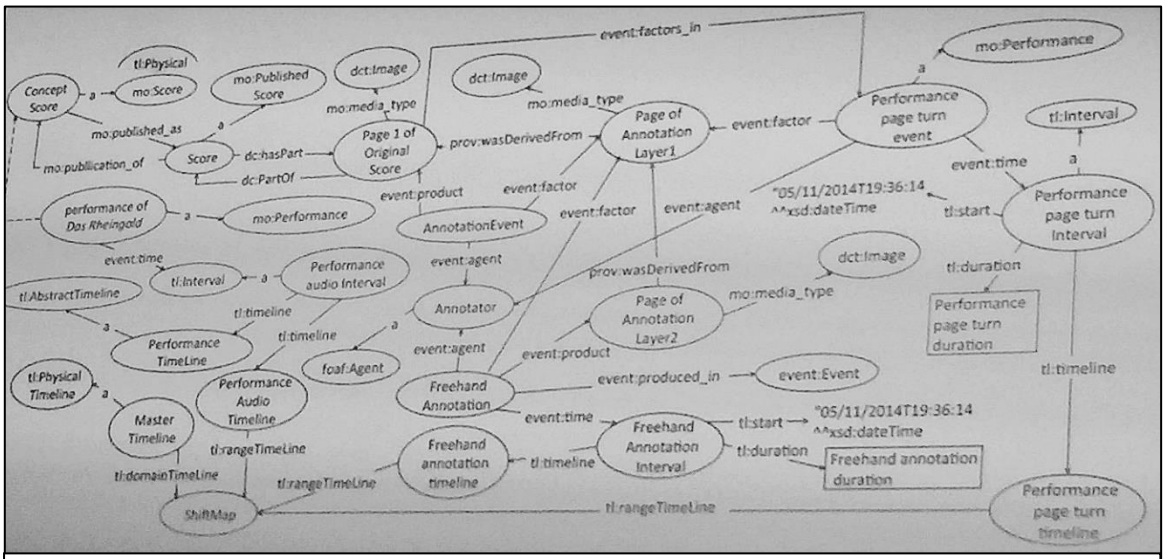

*Fig.2.3.*Graph of Wagner's Ring Cycle. (Page, 2017).

ambiguity (Cagnazzo, 2019), for example, making it clear the Beethoven in the above example is a composer and not a dog. While the subject and predicate always use URIs, the object can also just be a value (Shotton, 2013), such as, Franz Schubert → has middle name → Peter. It wouldn't make sense to link to an external 'Peter', so the name can be entered without a URI.

LD relationships are often visualised in a mind-map format, called "Graphs" (Sporny, 2012). Pennington (2019, p.157) states that these Graphs "model how all sources of shared information can and should be connected." Fig.2.1. demonstrates Pennington's (2019) dementia ontology model. It outlines the 3-step relationships in Graph form, but in human-understandable language. Syn (2019) demonstrates how to transition from the human-focused Concept Map into machine-readable language (Fig.2.2.) using ontology namespaces and hyperlinks for URIs. Page's (2017) Graph (Fig.2.3.) demonstrates some connections that could describe a resource with LD, although to be machine-readable the input would be written in XML code (Shotton, 2013). This informed my LD models in Chapter 5 and is discussed in more detail there.

Even if coding is a final step, visualising the Graphs helps humans see the relationships more clearly; mind-maps are an intuitive way to present knowledge. In fact, a project to preserve indigenous Alaskan knowledge used Concept Maps (Fig.2.4.) that presented the Subject → Predicate → Object structure in exactly this way (McCann, Pulsifer and Behe, 2016). Although the project didn't use LD, it could easily be added later.

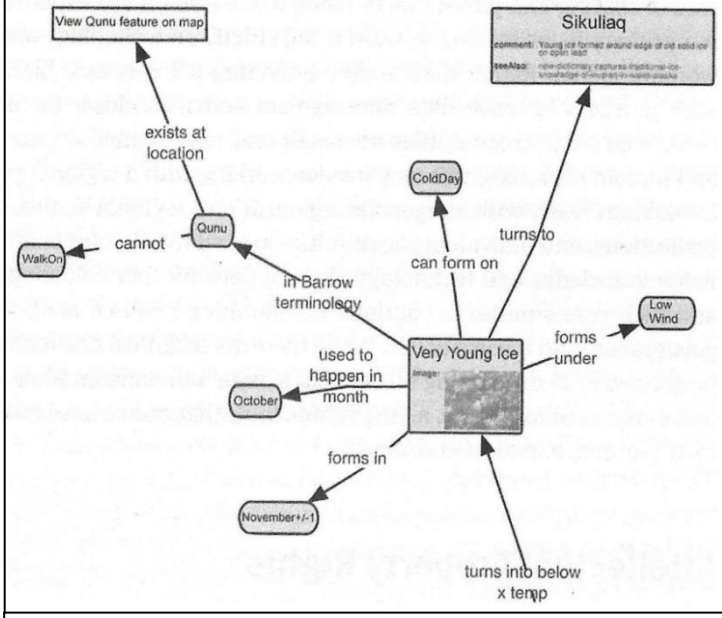

Fig.2.4. Concept Map created by indigenous Alaskans to describe ice (McCann, Pulsifer and Behe, 2016).

2.2.Best Practice

Van Hooland and Verborgh (2014) caution that LD's lack of a well-defined technological definition has allowed some technology sellers to disingenuously promote their product as supporting 'LD' if links to external resources are present at all. As mentioned above, LD should facilitate the Semantic Web, making it possible for computers to understand semantic relationship between items. Berners-Lee (2009) defines four LD principles:

1. Use URIs as names for things
2. Use HTTP URIs so that people can look up those names
3. When someone looks up a URI, provide useful information, using the standards (RDF, SPARQL)
4. Include links to other URIs, so that they can discover more things

Essentially, these principles specify that for something to be LD, it must use URIs, HTTP and RDF to connect related data online (Cagnazzo, 2019). However, other definitions exist. Shotton (2013) states that "anything published on the web in RDF *is* 'Linked Data'," completely ignoring Principle 4 about linking to external resources. Meanwhile, Berners-Lee's (2009) own five-star Linked Open Data system ranges from merely making resources openly available on the web (one star) to using an open, machine-readable, untrademarked format with RDF to identify resources and links to other people's data to provide context (five stars). The external linking is only required for a five-star rating, and RDF appears at 4 stars. Although this was created as a development scheme, and not an LD definition (Hausenblas, 2012), Berners-Lee provides a star just for information being online, demonstrating the wide technological discrepancy between institutions, as well as the range of things operating under the guise of LD. This is a problem since, as Syn (2019, p.55) argues, LD makes information retrieval and discovery better for users, but only if data uses "the right format and the right linkages to other data points."

Fig.2.5. You can buy cups featuring Berners-Lee's 5-star development scheme. (Berners-Lee, 2009).

Pennington and Cagnazzo (2019, p.19) make recommendations for Scottish libraries. These include having a clear strategy and making a start with LD "even on something small." Engagement and collaboration with others are important, with experts such as universities and outside developers but also the wider community. O'Dell (2016) agrees with them that controlled vocabularies are useful for maintaining consistency across LD.

2.3.Linked Data for Libraries and Small Archives

Libraries, museums and other cultural heritage bodies have often run into the problem of having many fantastic resources that are stored away, and split by field, in not-readily-accessible buildings

without many indications to the general public that the resources exist at all (Thorsen and Pattuelli, 2016). Library catalogues have traditionally only provided information about items in the library's own collection (Spiteri, 2019). Although union catalogues, such as WorldCat, allow libraries to upload and share their records in one place, making it easier for other libraries to catalogue their own collection and for users to search for resources across libraries (What is WorldCat?, 2019), they're just larger silos. Since they're not fully intertwined with the rest of the web, users need to seek out the catalogue to search it. Coyle (2010, p.5) acknowledges that, although union catalogues bring together library collections, they are still "an information environment separate from the web." As Pennington (2016, p.36) argues, since people are more likely to search for information online using Google or Wikipedia, "if libraries do not ensure that their data is opened up to be findable and retrievable on common websites, the world will miss the expansive resource that they offer." Coyle (2010, p.5) argues that library catalogues need to shift from being merely "on the web" to "of the web", pointing to LD as a tool that can help libraries become more interconnected with other online resources. Fay and Sauers (2012) agree that LD allows collections to be more easily searchable and gives the patron greater control to filter information, making it easier to find relevant resources. Spiteri (2019) adds that LD can also allow library catalogues to be a much richer resource by providing access to additional related material outside the library's collection. So, the main advantages of LD for libraries are that it connects outside searchers to the catalogue and catalogue browsers to related items of interest outside the collection.

Some libraries, like the British Library, already use LD, (Hill, 2017). Libraries and other cultural institutions must adapt to LD because the Semantic Web is already becoming a reality and libraries will need to modify their catalogues if they are to remain relevant. Hogg (2015, p.55) states that with more digital library resources, "the challenge for smaller libraries and 'niche' collections will be to ensure that our presence is recognised and discoverable among the Goliaths of information providers." If libraries are to adapt to LD, Pennington (2016) emphasises that the RDF standards must be implemented consistently for the system to work as intended. Fay and Sauers (2012) acknowledge the process of crossing over to an LD system requires special provision of time, funding and staff input, which means that smaller archives face the challenge of having less access to the resources that larger "authorised" libraries are able to draw upon (Roberts and Cohen, 2013, p.243). Financial challenges limit smaller cultural heritage institutions in digital enterprises (Hogg, 2015); nonetheless, as Adrian Stevenson in Ruddock and Stevenson (2011, p.19) argues, LD can help expose the "hidden" collections of archives and special collections, making them more used.

2.4.Linked Data and Music

The International Association of Music Librarian's (IAML) Cecelia database of music resources across the UK and Ireland demonstrates that individual collections' music materials are already "seen [by music libraries] as part of a wider shared collected resource" (Andrews, 2003, p.202). This early attempt to break down silos indicates that LD could be valuable for Music Libraries since it could help further this goal.

Music cataloguing has often been at odds with traditional text-based cataloguing, due to extra requirements from patrons and cataloguers. Papakhian (2000, p.21) complains that music cataloguing plays "second fiddle in the library world", as does the bibliographic control of music and music materials. This is an old argument; Redfern (1979) identifies the difficulty for music cataloguers in standardising their collection catalogues because the same music often has different titles depending on publication, more than traditional literary works. Older cataloguing standards guided cataloguers to record the title of an item exactly as it appeared on the title page, making retrieval of the item more difficult for both patrons and librarians (Redfern, 1979). With LD, URIs act as uniform titles, joining manifestations of a work together and aiding with resource retrieval.

The prevalence of music-specific LD datasets observed by Gracy, Zeng and Sirvin (2013) and corroborated by McCrae et al (2019), who identifies thirteen creative-commons licensed resources with music-specific LD datasets, indicates that cataloguing might also be more balanced with LD. Raimond et al's (2013) Music Ontology provides 54 classes and 153 properties for use on the Semantic Web, each with its own URI, and, so far, it has been widely adopted (Gracy, Zeng and Sirvin, 2013). Page *et al*. (2017), for instance, enhanced an online Live Music Archive with LD using the Music Ontology alongside other non-musical ontologies.

Fay and Sauers (2012) acknowledge the challenge of searching for audio files on the web, since there are few central storage sites, and generally the metadata on these files is not easily searchable if, indeed, it exists at all. LD offers a way to make these audio files searchable, by using RDF predicates to supply the computer with more information about these non-textual files. Rose-Steel and Turnator (2016) found the freedom to create many connections with LD useful when cataloguing medieval motets, but this extends to music cataloguing generally because of the ability to easily link to information about, for example, key, tempo and instrumentation (Raimond *et al*., 2013). The wide variety of media held by music collections may have historically challenged music libraries (Pugin, 2015) but perhaps this now presents extra appeal for digital musicologists because a single music collection can provide many avenues to explore with LD (*Transforming Musicology*, 2018).

2.5.Librarians and Tech

Dobson (2019) identifies a problem with traditional search and retrieval systems – they are difficult to use, and probably require training. Semantic web requirements can also be intimidating and discourage participation (Cagnazzo, 2019). Despite ever-increasing technology requirements, Batthini (2014) highlights that many librarians' computer skills are self-taught, although library courses teach up-to-date tech skills (Riley-Huff and Rholes, 2011). Recent attempts to create Linked-Data social hashtags are still in the trial stage (Syn, 2019), but are appealing because they combine the user-friendly, social media focus of Web 2.0 with the disambiguated and structured approach of Web 3.0 – Cagnazzo (2019) calls this the Social Semantic Web. Using popular hashtags has already been shown by Chang and Iyer (2012) to be an effective way to link users to library material about the hashtag's subject. LD hashtags would make tags compatible across systems, increasing the effectiveness of information discoverability, so could be a useful tool for libraries and other cultural heritage institutions. Bates (2013, p.80) argues that "a collection of information becomes more valuable when its users can engage with that information", and Spiteri (2019) suggests that this reasoning can extend to library discovery systems. Although Pennington and Cagnazzo (2019) found that some librarians are already interested in learning about LD, "more user-friendly interfaces" would make LD easier to engage with (Cagnazzo, 2019, p.33).

2.6.World Music Ethics and Linked Data's Collaborative Possibilities

Roy (2016) explains that the Western values of Open Access and freedom of expression can come into conflict with traditional cultural taboos about when and by whom some content can be accessed. The British Library has a World Music collection online and mentions the ethics of sharing "ethnographic sound recordings" (The British Library Board, 2009), which are pertinent for all World Music collections. The British Library Board (2009) stresses the rights of the "intangible cultural heritage" of the music originators, and the ethics of sharing music from indigenous cultures, indicating that non-profit sharing of World Music for preservation and research purposes is ethical, while commercial exploitation is not. Francis *et al.* (2016) and Pasaribu (2016) both suggest that password-protecting sensitive material is a good compromise between providing access and protecting cultural taboos.

Brandt (2016, p.221) acknowledges the historical power imbalance of privileged academics and ethnologists over the indigenous people they studied, photographed, filmed etc. but also points to positive aspects that are emerging as indigenous people "transition from oral tradition into fully literate society." Indigenous material protected in museums and other cultural heritage institutions helps to enrich indigenous societies' links to their past. Some indigenous communities are more

enabled to collaborate and network with their collections than they have been in the past (Chisita, Rusero and Shoko, 2016), and Dempsey (2000) notes that the rise of the internet has made this possible. Indigenous community-run cultural heritage projects are increasingly emerging (Villanueva, 2016; Chen, 2016) and Villanueva (2016) agrees that the internet makes it easier for indigenous libraries to access materials that were taken during colonialism. Francis *et al.* (2016) mention that indigenous sound collections are often under-catalogued, since archivists rarely have the time or resources to catalogue indigenous audio materials deeply. Even if they do, Francis *et al.* (2016, p.347) argue that Western archivists cannot have the "knowledge necessary to include details and categories salient to the source communities", stating that the only way to overcome this is for the archive and source community to work together.

Spiteri (2019) identifies the knowledge and expertise of library users as a valuable resource for LD. She goes on to mention that users can be guided in their participation in the creation of LD system by providing predefined fields (facets) for users to fill in. "I think we need to be careful of dismissing the contributions of our users as being that of amateurs, especially if they could, in fact, know far more about the topic than library staff" (Spiteri, 2019, p.105). While the Fanshawe Archive, particularly FM, is very knowledgeable about the collection, LD may facilitate contribution from scholars or people from the recorded cultures. This provides deeper insight into the collection, while empowering the source communities to contribute in sharing and describing their culture.

CHAPTER 3 – METHODOLOGY

This project comprised of a case study with mostly qualitative research methods. This provided a microcosm to explore challenges, opportunities and practicalities that must be considered when developing an LD catalogue, through the lens of a small archive with limited resources. The qualitative approach is generally interpretivist (Thomas, 2011), exploring how the archivists understand their role and the significance and vision of their archive. Conducting interviews and documentary analysis is therefore appropriate for gaining a deep understanding of the Archive's processes (what they do) and systems (the way they organise what they do). This then informed how best LD might be implemented.

3.1.Literature Review

A literature review of best practices, current models etc. for creating LD, particularly in a music context, provided an understanding of the fields of LD and music archives and informed the refinement of research questions. It also supported recommendations for how the Fanshawe Archive might work towards implementing an LD cataloguing schema.

3.2.Project Research

The main research was conducted through interviews with the Fanshawe Archive (internal interviews) and other libraries (external interviews) to place the Archive in its wider context. I also conducted a week-long site visit, to better understand the Archive's specific layout, situation and resources. The internal interviews were all conducted in person during the site visit.

3.2.1.Interviews

I conducted internal interviews with the Fanshawe Archive, and external interviews with other libraries (three small music libraries, and a larger national library that requested to remain anonymous). The internal interview questions centred on the Archive's vision, processes, resources, schedule and target audience, as well as discussing the Archive's current digitisation and cataloguing projects and the possibilities of using LD to increase its accessibility. These were very useful for examining the Fanshawe Archive's individual situation and gaining an understanding of what the Archive consider valuable and useful. They helped inform suggestions specifically for this archive.

The external interviews were also extremely useful for placing the Fanshawe Archive's situation in its wider context. Although the anonymous national library (ANL) was the only library that currently uses LD, how the smaller libraries were tackling online challenges provided ideas for the Fanshawe Archive's online presence. They also provided further insight into the limitations and successes for

smaller institutions, as well as information about the wider understanding and interpretation of LD. While the interviews mostly involved qualitatively recording the archivists' and librarians' unique interpretative worldview, some quantitative aspects were also covered, such as the number of resources in each format, the time it takes to catalogue or digitise a single resource, the general userbase size etc.

The interviews were with professionals about their work, so while they had to be framed in a safe context where the interviewees could, for example, feel free to criticise LD, the level of sensitivity was less than it would have been if interviewing members of the public or disadvantaged social groups as noted by Richards (2006). The interviews were semi-structured, which Thomas (2011, p.164) describes as the "best of both worlds" between structured and unstructured; the list of interview topics can be combined with freedom to explore points that arise in more detail. I chose this approach because it would make the interviewees the most comfortable. The semi-structured format provided guidance for the interviewees, so they knew what I was asking of them, but they were still able (and encouraged) to elaborate if it seemed relevant. By going down suitable tangents, often the interviewees provided information that I wouldn't have known to ask about. For example, although I knew to ask about cataloguing, my personal cataloguing experience is limited so I wasn't aware of the significant difference between cataloguing and indexing, but indexing also came up a lot in the external interviews and that informed some of my suggestions for how the Fanshawe Archive might present the Fanshawe Journals.

	FM	FPT	FT(D)	FT(L)	GCH	VWM	ANL	PCL
Linked Data Understanding	Y	Y	Y	Y	Y	Y	Y	Y
Vision for Institution	Y	Y	N	Y	Y	Y	Y	Y
Patron Questions	Y	Y	N	Y	Y	Y	Y	Y
Making Institution Accessible	Y	N	Y	Y	Y	Y	Y	Y
Previous Projects	N	N	N	N	N	Y	Y	N
Resource Questions	Y	Y	N	N	Y	Y	Y	Y
Current Tech Situations	Y	Y	Y	Y	Y	N	N	Y
Technical Requirements	N	N	Y	N	N	Y	Y	N
Catalogue	Y	Y	N	Y	Y	Y	Y	Y
Getting Support	Y	N	N	Y	Y	Y	N	Y
Giving Support	Y	N	N	Y	Y	Y	Y	Y
Ethics of using Indigenous Sources	Y	N	N	N	N	N	Y	N
Future Hopes for LD	Y	Y	Y	Y	N	N	N	N
Does Interviewee have any other comments?	Y	Y	Y	Y	Y	Y	Y	Y

Fig.3.1. Interview Theme Chart

I numbered the interview questions to make analysis easier and to keep a consistent approach across all interviews. I created a theme chart, to decide which areas to ask each interviewee about

(Fig.3.1.). Each theme was labelled by letter, and these are the same across all interviews even when an entire section was not included for a particular interviewee.

However, within each section, the subheadings and questions were numbered consecutively according to that specific interview, meaning the related questions wouldn't have the same number across interviews but were within the same alphabetical section. This structure was chosen to keep the overall sections distinct for purposes of comparison while making the Interview Questions Script easily navigable for the interviewees, who saw questions in advance.

The interview questions sent to the participants are included as an appendix. I considered attaching the full interview transcripts as appendices, while giving participants the option in the consent form (Appendix A) to opt out. However, the ethics committee suggested "not including the transcripts as appendices (even if you allow the participants to ask for this not to happen). I suspect this whole idea might put them off."

I chose to keep the questions broad (not LD focused), particularly in the technology and support sections, because the interviewees' institutions were not necessarily undertaking LD projects themselves. Even outside LD, these questions determined what resources the smaller libraries and archive had access to and where they would go for support for large-scale projects (like the LD one would be). This helped me ascertain how feasible implementing LD would be for the Fanshawe Archive. I made sure to include the opportunity to explore the potential drawbacks that the Fanshawe Archive would experience if they adopted an LD readable approach for their catalogue.

For the internal interviews, it made sense to get as many perspectives from the Archive as possible, and I was able to interview the Fanshawe Manager (FM), who was providing me with accommodation, and the part-time member of staff (FPT) who lived nearby. Happily, my site visit coincided with an archive filming event that both Trustees (FT(L) and FT(D)) had travelled over to be involved in, giving me the opportunity to interview them both in person. I realise that using letter abbreviations for the interviews can be confusing, so all the internal interviews begin with an F for Fanshawe, whereas the external interviews start with different letters.

I contacted the anonymous national library interviewee (ANL) because I had attended their talk about the library's recent LD project and thought they would be able to provide the valuable perspective of a large library already working with LD. The smaller libraries were chosen to get a sense of how music libraries of a similar size to the Fanshawe Archive were functioning with regards to online cataloguing, digitisation and making their resources available online, even if they weren't using LD specifically. I had attended the Vaughan Williams Memorial Library's (VWM) presentation about their recent project (The Full English) to make English Folksongs more accessible online to a

wider audience, and since the literature often indicated that LD could help raise a library collection's online profile, I felt it would be useful to talk to the non-classical music library about the project, even although it didn't use LD. The Piping Centre Library (PCL) also deals in traditional non-classical music, is staffed by a lone librarian, and their online catalogue, while further along than the Fanshawe Archive's, is still a work in progress. Their similarities with the Fanshawe Archive made them a valuable participant. The Gerald Coke Handel librarian (GCH) had written an article about getting their catalogue online, specifically about digital challenges for small libraries, which was also useful.

3.2.2.Site Visit

Although the interviews were my main source of research data, the site visit to the Archive was extremely useful. As well as allowing me to conduct the interviews in person, I saw how the Archive works and the items within the Archive. It wasn't until I visited the Archive that I really understood the layout and how the items fit together. The Fanshawe Manager and I walked through the process of collecting the different materials that related to a few particular recordings. The Archive is laid out logically, but there is still currently a complex process of getting the date from the Masterbox, going to the journals and looking through them for the entry that correspondes to it, then hunting though the slides for the corresponding place and date. The Fanshawe Manager had sent me some material prior to the visit, including an outline of their ideas for an online catalogue and a sample of their current catalogue, but I struggled to understand it until I saw in person how the catalogue corresponded to the Archive's materials. At the Archive, I was given the chance to listen to digitised versions of the tapes, watch Fanshawe's documentaries, listen to radio interviews with and about Fanshawe, and read the journals. This documentary analysis showed me the wide range of materials that the Archive holds and gave me access to some materials I couldn't have accessed anywhere else. This helped inform the creation of LD Concept Maps.

Since the Fanshawe Archive is a very personal collection, looked after by people with a close relationship to the Archive, it was also helpful that I got to witness that. Because I was so immersed during the visit, I created Fieldnotes while I was there to help me remember and process the information I was absorbing through observation, documentary analysis and conversations with the Fanshawe Manager and others.

3.3.Linked Data Model

I acquired a small sample of already-digitised materials from the Archive to demonstrate how LD could be implemented to suit the Archive's specific context. One of the biggest challenges with changing a catalogue over to LD is getting started with a model. This project offers suggestions and

outlines the process of deciding on an ontology and demonstrates how the Archive might begin to implement LD. Since each sound recording in the Fanshawe collection has so much related material, I created two examples: one that looked at how the Archive could link related materials within its collection, and one that looked outwards to how the Archive could link to relevant external resources. To vary the examples, one was a single instrument, played in a traditional style, from the African Collection and the other was a boundary-pushing band that used improvised instruments from the Pacific Collection. Although full examples weren't given, I also discussed other items from the collection that posed their own considerations, including vocal music, and obscure instruments that might be difficult to link outward. I used the literature review to inform how to go about creating the LD models.

Whether the Fanshawe Archive chooses to adopt LD or not, these models could help guide others, since they focus on the kinds of links that could be made, through the lens of the Fanshawe collection. Limited resources, particularly time, impacted on how detailed the examples could be, but the purpose of the examples is to demonstrate how to look at LD in the context of the Archive, not to fully catalogue the Archive, so the limited extent is reasonable for the purposes of this dissertation. It would be up to the Fanshawe Archive to decide the best links to create, and where to stop with their own LD project.

3.4.Further Discussion

This methodology supported the examination of the Fanshawe Archive's specific context, looking at the challenges, feasibility and opportunities posed by LD, while also considering the wider state of LD and the Semantic Web to judge whether the LD system is worthwhile for small archives.

The background literature research, site visit, internal and external interviews, and discussion and creation of example starting points, allow me to recommend best practices for LD within small archives and provide suggestions for how to tackle the cataloguing process – regardless of whether it is concluded that the LD process is appropriate for this particular case.

CHAPTER 4 – ANALYSIS 1: THE ROUTE FROM ANALOGUE TOWARDS LINKED DATA

Chapter 1 describes the two projects the Fanshawe Archive is working on: creating a catalogue and digitising the collection. It is reasonable that, since the Archive's attention is focused on these projects, any considerations about LD are a minor priority. This chapter therefore brings together the research from the site visit to the Archive and interviews conducted with the Archive and other libraries to examine how the Archive is carrying out the projects in the context of other libraries' practice. In both cases, I explore options the Fanshawe Archive might take now to make LD easier if they choose to implement it in future. Since small heritage institutions like the Fanshawe Archive often lack the resources to carry out such large projects alone, collaboration and support are discussed. Next the ethical considerations that arise when Western institutions make World Music available, particularly online, are examined. Finally, the chapter looks at the benefits and challenges of incorporating LD for the Fanshawe Archive and similar small archives and libraries.

4.1.Creating a Catalogue

The Fanshawe Archive wishes to create a catalogue to help users browse and search their collection. As mentioned in Chapter 1, they plan to transfer the catalogue to an Excel spreadsheet, but with a more user-friendly interface display (FPT, 2019). According to Sporny (2012), Excel can work as an initial way of visualising LD because it corresponds to the RDF format of Subject → Predicate → Object. This is demonstrated further in the recommendations. Having a user-friendly interface for their catalogue will probably require subscribing to or purchasing software. Various online catalogue software providers exist, and each external library interviewed uses different catalogue software. The anonymous national library's sound catalogue software is not the same as for its book catalogue (ANL, 2019). Different software has different benefits. The Vaughan Williams Memorial Library (VWM, 2019) uses the archival CALM software to structure hierarchically: allowing them to arrange by collection at the highest level and then break each collection down in various ways. The Fanshawe Archive might find such a hierarchy useful to structure Fanshawe's different journeys.

The Gerald Coke Handel Library (GCH, 2019) inherited a private collection with a card catalogue and files, so were recently in a similar analogue position to the Fanshawe Archive. Now they have an online catalogue through the software provider Soutron LMS, an external company that installs and maintains the portals for an annual charge. Their catalogue has also recently been uploaded to the National Bibliographic Knowledgebase (NBK), a union catalogue of over 150 institutions with hopes to reach over 200 in 2020 (Ruddock, 2019). The Piping Centre Library's (PCL, 2019) catalogue is also hosted on a union catalogue, Opals, which they find useful because, although their collection is specialised, they can save time and resources by downloading some catalogue records from other

libraries. PCL (2019) told me that union catalogues are perhaps a step down from LD, but I feel the prevalence of union catalogues demonstrates that libraries are trying to be interconnected despite limited resources. For now, union catalogues are a more feasible option for smaller institutions than an LD scheme, especially since, as ANL (2019) mentioned, there are currently no off-the-shelf LD programs for libraries.

4.1.1.Detailed Online Catalogues and Deep Indexing

One challenge for cataloguers, as FT(L) (2019) mentioned, is knowing where to stop. GCH (2019) described trying to make their online catalogue as detailed as possible. Many foreign patrons visit to view a special item, so it is important that they can access accurate information before making the journey to ensure they don't waste a trip because they've mistaken the song the collection carries for a different one. This is relevant to the Fanshawe Archive since it is likely that many online users will be from Africa or the Pacific, especially since the Archive has ties to the University of Fiji and National Film and Sound Archive in Australia (FM, 2019). Since the Archive might decide not to put any culturally sensitive recordings online, it will be necessary to ensure that those items are still catalogued accurately and deeply. They certainly don't want a researcher from Fiji to make a special visit to hear a sensitive piece of music only to discover that the recording is not actually what they were looking for. The collaborative nature of LD, discussed further in Chapter 5, could help ensure the Fanshawe catalogue is accurate.

Interestingly, the Vaughan Williams Memorial Library decided to catalogue their collection deeply, down to individual songs on a ballad sheet (VWM, 2019). The ANL, meanwhile, just completed an LD project where they opted to group some items together in a single record. ANL (2019) explained that for groups of miscellaneous sketches, which are physically kept together in a folder, they decided that it was not worth the time-consuming and complicated task of splitting up the digitised item. Nevertheless, they talked about the "tension" of trying to replicate a user's experience with a physical item as closely as possible with their digital representations. ANL told me the library considered creating separate digital records for each sketch but explained that, particularly in cases that would require extensive recataloguing, "it's better to preserve and replicate the effect the user would have looking at the original than kind of splitting things off and giving some skewed view of what it's like in reality." VWM, on the other hand, described how their CALM software's hierarchical structure meant that they could catalogue down to the individual song, and avoid records such as "a box of correspondence", which was more useful here because the songs were the focus of their collections. Clearly both approaches are valid, and the decision mostly comes down to how

worthwhile separating the items would be for the end user. Sometimes splitting the items would be less beneficial than keeping them together.

The depth of the Fanshawe Archive's cataloguing is an important consideration. It would be worthwhile for each photograph and sound recording to have a separate record, particularly since those records could be used as URIs for LD. Within the journals themselves, however, the entry for any given sound recording is generally a few lines at most, so creating an individual record (or URI) for each entry would be impractical.

I would suggest creating a single catalogue entry per journal because they are more useful if they can be easily accessed in their full form. Users will be able to see the single entry in the context of others and get a better feel for Fanshawe's journeys and processes. A deep level of indexing would still keep the individual entry easily accessible, since the Fanshawe Archive could catalogue the journal and then index each journal entry to let the user know what and where sound recordings are discussed. Three of the four external libraries mentioned the importance of high-level indexing to help users get the most out of online catalogues. Detailed indexing can help users navigate the physical item (ANL, 2019) but can also help them determine from an online record whether an item contains the piece of music they're looking for (GCH, 2019). As mentioned above, many of the Archive's online users will be accessing the catalogue from a distance, so it is important to avoid misunderstandings in the descriptions.

4.1.2.Cataloguing Music and Vocabulary Control

Although Papakhian (2000) and Redfern (1979) describe music cataloguing as particularly challenging, the external interviews indicated that the music librarians didn't find cataloguing their resources difficult. The librarians at the Gerald Coke Handel Library have the music literacy and language skills to interpret the music they catalogue, and have added extra MARC fields for specifically musical things like rhythm number or thematic catalogue number (GCH, 2019). The Piping Centre library also uses MARC records, PCL (2019) explaining that cataloguing the music isn't especially complex: "a record's a record." PCL mentioned, however, that bagpipe cataloguers need to understand the important distinction between Light Music and Pibroch. Although Pibroch sheet music exists, it is tied to bagpiping's aural tradition and it is expected that a bagpiper will learn a piece of Pibroch with an instructor. Users will want to know the player, since interpretation is integral to Pibroch music. When cataloguing a record of solo bagpipe music that has a piece of Pibroch being performed, the Piping Centre librarian makes sure to include that information in the record, utilising a field for music-specific purposes in a similar way to the Gerald Coke Handel Library.

This raises the question of whether there is anything in the Fanshawe collection similarly relevant to specific World Music cultures that might be helpful for users to know about. The Fanshawe Manager and journals could perhaps identify some examples, but the collaborative nature of LD means that even if some things were missed, indigenous cultures could engage with the catalogue and help provide context through LD cataloguing.

One complication arising in both Western classical music and in folk or World Music is that the same piece of music is often referred to by several titles. In the Western Classical tradition, for example, Beethoven's Symphony No. 3 can be Beethoven's Third, Heroic Symphony, Eroica Symphony, older writing might refer to it as Buonaparte, etc. (ANL, 2019; GCH, 2019). Similarly, folksongs often have multiple titles. The folktune 'Prickerly Bush' has many spelling variations but is also known as 'The Maid Freed From the Gallows' (VWM, 2019). Western classical music uses opus or thematic catalogue numbers to identify composers' works (GCH, 2019), the Vaughan Williams Memorial Library turns to the Roud Folksong Index for streamlining English folksongs (VWM, 2019), and the Piping Centre Library uses Bob Pekaar's Tune Encyclopaedia to identify bagpipe music (PCL, 2019). LD uses URIs which can bring together items that are the same. Using the FRBR model, if these indexes were available as LD datasets, the library catalogue entries could link to the URI for the work, joining together manifestations with different titles.

The Gerald Coke Handel Library – a collection of Western classical music – uses uniform titles, authority-controlled fields, and thesauri to describe their resources (GCH, 2019). So far, the Fanshawe Archive have been unable to find suitable external vocabulary controls for their collection (FM, 2019). They plan to create their own thesaurus for their Excel catalogue with dropdown picklists for as many fields as possible (FM, 2019). One advantage to this approach is that it might enable a volunteer cataloguer who was less familiar with the collection to catalogue to an appropriate standard. FT(L) (2019), who worked as a librarian, had concerns around volunteer cataloguers: the catalogue is such an important access portal to the collection, it has to be arranged consistently, and it has to be user-friendly to someone not familiar with the collection. "I don't have a problem with volunteers" she told me, "but it's using and working with volunteers and giving them appropriate tasks that are rewarding for them but actually can be achieved to the standard that you need… you're not helping the volunteers by giving them something where they can't produce a good result because they don't have the skills and the training." If the Archive chose to take on a volunteer student then a picklist could help guide them. A picklist could also be compatible with LD; it resembles Spiteri's (2019) suggestion that LD creators could be steered by cataloguers providing pre-defined facets to fill in.

A potential danger of a small institution like the Fanshawe Archive creating their own thesaurus is that, even if their collection is internally consistent, it should also be consistent with other online resources, particularly if they might take an LD approach in future. While it would be possible for them to change their vocabulary when they switched to LD, this needlessly doubles the work. If they model their thesaurus to align with LD now, it will be easier for them to switch to LD later. So, the Archive needs to be careful to choose vocabulary that is compatible with both external and internal resources.

4.1.3.Creating a Linked-Data-Ready Catalogue

Although the Fanshawe Archive is unlikely to move from a Microsoft Word File directly to an LD structure, there are certain choices they can make that will make using LD easier in future. For LD, it makes sense for each catalogue record to act as the URI for the item being described, so each sound recording will have a catalogue record that represents the sound recording itself. Looking at the catalogue software providers used by the external libraries, it became clear that some online catalogue software lacks permanent links for each catalogue record and copying the record's web address returns a Page Not Found error after the internet window is closed. I would recommend, therefore, that the Fanshawe Archive choose an online interface provider that uses readily accessible permalinks that can later be adapted into the URI for each item.

The lack of any off-the-shelf LD software is a challenge but, taking an idea from Sporny (2012), the Archive can still use the RDF format (without the Machine-Readable coding) in their Excel spreadsheet using the rows as the Subject, the column headings as the Predicate and the value as the Object (Fig.4.1.).

	A	B	C	D	E	F
1		Recorded in	Recorded by	has Medium	has Number of Performers	
2	FPC 88 Gilo Stones	Solomon Islands	David Fanshawe	Magnetic Tape	2	
3						

Fig.4.1. RDF Format in Excel

The format could then make it easier for the Archive to create RDF triples in future.

If a full LD project isn't feasible, it might still be possible for the Archive to include some columns in their spreadsheet that link outwards as hyperlinks to external resources. The Piping Centre librarian (PCL, 2019) hypothesised that including hyperlinks in MARC fields might be a compromise, linking to external resources at a document level. If the Fanshawe Archive did this in their catalogue, then when they come to create the LD, they will have already identified useful external resources that they can link to.

Finally, as mentioned above, for their catalogue to be compatible with other collections, the Archive should consider using structures and controlled vocabulary that align with external standards as much as possible, even if in some cases, they'll need to come up with terms themselves.

4.2.Digitisation Project

There is a gap between library institutions. While much library material is available online, many cultural institutions, small ones especially, still have a lot of material exclusively in analogue format. The Fanshawe Archive is one of many small institutions trying to get their collections into a more accessible digital format. The Vaughan Williams Memorial Library paid to digitise some of the collections used in their online folk music project, the Full English (VWM, 2019), and the Piping Centre Librarian (2019) mentioned that they're "a bit behind the curve" with digitising their old sound materials. Unlike the Fanshawe Archive, much of the Piping Centre Library's material is not unique, but some of the rarer recordings are fragile, particularly the wax records from before 1960 and reel-to-reel tapes.

4.2.1.Advantages and Necessity of Digitisation

The Fanshawe Archive initially set out to digitise their collection as a preservation project. Some of the 'Sticky Tapes' are deteriorating faster than expected, and the slides risk decolouration (FM, 2019), so the Archive is making copies so they don't lose anything if "something goes pear-shaped" (FPT, 2019). Although the topic of 'sticky tapes' wasn't raised in the Piping Centre Librarian's interview, he corroborated that reel-to-reel tapes degrade after a while (PCL, 2019). But, beyond this, the Archive recognises that digitisation has benefits outside preservation and digitisation is an important step towards making the collection accessible to a wider audience online (FPT, 2019).

One of the Fanshawe trustees, FT(D), is a skilled sound engineer, so is working towards restoring the most at-risk tapes. As one person working to digitise 3,200 tapes, the process is slow, but an advantage of the Fanshawe collection is that it is finite and will not be added to. This means that once the collection is digitised, that part of the work will never need to be revisited. This contrasts with the external libraries interviewed, whose collections are still growing (ANL, 2019; GCH, 2019; PCL, 2019; VWM, 2019).

FT(D) explained that, provided the files are saved in the right format, digitised items are much easier to copy. Multiple copies can be stored in different places, so the collection is safer. However, the digitised files need to be stored as WAV files so they didn't lose quality over multiple copies (which they would if stored as compressed MP3s), and WAV files take up considerably more space than MP3s. We tested the tape that FT(D) was digitising at the time, and there was already a big

difference size for the single file: 4,933KB for the MP3, compared to 118,043KB for the WAV file. They have 40,000 items, so the amount of storage space the archive would need to pay for is going to skyrocket. Compression ruins the quality of the recording, and since Fanshawe took such care, it would be terrible to lower their quality now, purely to save space.

The Fanshawe Archive are keen to make the journals available digitally online, albeit with some personal content removed (FT(D), 2019; FT(L), 2019). GCH (2019) described their involvement with RISM UK, where they uploaded some manuscripts to a program that mimics turning pages. Financial and time resources are an important consideration (RISM UK paid the Gerald Coke Handel Library and others to be involved in the project) but I feel that an application like this would be a fantastic way for the Fanshawe Archive to display the journals, with any sensitive material removed, of course. As discussed in the cataloguing section, ANL (2019) mentioned the value of making digitised objects online replicate the effect the user would have looking at the original, and software like that offered by RISM UK would help the Archive live up to that.

The Archive is about to be involved with the Heritage Lottery-funded Save Our Sounds project, which will digitise 600 tapes (FPT, 2019), helping speed up the Archive's digitisation process (FT(L), 2019). Save our Sounds was established to preserve at-risk sound recordings given that so many small archives are in a similar position, with sound collections in danger of physical degradation (*Save Our Sounds*, 2019). The project will only digitise the tapes, not the slides, journals or other ephemera, but this will still significantly boost the Fanshawe Archive's digitisation progress. The Archive didn't take the decision to get involved with Save Our Sounds lightly; they want the preservation process to be done properly. The Fanshawe Manager and a trustee (FT(D)) visited the project to see the digitisation process (FT(L), 2019), and are satisfied with the quality – although they still plan for their own expert, FT(D), to digitise any tapes that have the 'sticky tape' complication (FM, 2019). In addition, FT(L) ensured that the cataloguing side of Save Our Sounds was managed to a satisfactory standard; "there are volunteers involved," she told me, "but we're quite clear that they must not be handling some of this material or cataloguing it" (FT(L), 2019). This reinforced to me the care that the people at the Fanshawe Archive show for the collection. It also demonstrated that, although a private collection, the Trustees have the industry knowledge to properly vet the projects they get involved with to ensure their vulnerable recordings are treated with the necessary skill to keep them safe.

4.2.2.Digitisation and Linked Data

After the Save Our Sounds project has completed their input, the African Sound Recordings will be fully digitised (FPT, 2019). This huge milestone means that the Archive will have a discrete, fully

digitised collection that they can start looking towards the next step with, even as they continue to complete the much bigger Pacific Collection. They don't need to finish digitising the full archive before they can start putting their collection online, and while it is important for the physical collection to remain together, both trustees (FT(D) and FT(L)) acknowledged that their online presence could not include the entire archive immediately. I think the Archive would benefit from taking inspiration from the Vaughan Williams Memorial Library's online projects. The Vaughan Williams Memorial Library started small with a project called Take 6, where the library digitally preserved and presented online six of their folksong collections, before progressing to the much bigger Full English project (VWM, 2019). As well as being "a perfectly good thing in itself", the smaller Take 6 project helped to pilot the more ambitious Full English. It demonstrated the library's ability to handle the kind of project they were proposing with the Full English and helped them in their bid for Heritage Lottery Funding (VWM, 2019).

The Vaughan Williams Memorial Library don't use LD, but their methods could inform how the Fanshawe Archive might approach an LD project. If the Archive were to start small, either with the African Collection or perhaps even smaller, focusing on just one country, the Archive could start the learning process of using LD to place their small sample into a wider context. The small sample size would make the project more manageable and less daunting, giving the Archive the freedom to really familiarise themselves with one country's tradition and make the LD connections as detailed and useful as they deserve to be. Perhaps there is scope for a separate Masters Dissertation or PhD to take a small sample from the Africa Collection, once it is digitised, and investigate creating an LD system that presents the material in a user-friendly and informative way.

The Vaughan Williams Memorial Library were selective with the material they used for their project; "the idea was to have the most important collections online" VWM (2019). Similarly, the Fanshawe Archive should consider which parts of their collection would be most interesting to outside users: perhaps the recordings of traditions that have been wiped out by oppressive regimes, such as the Marsh Arabs (not African) and Ugandan Bwala Dancers recordings, targeted by Saddam Hussein and Idi Amin, respectively (Sweeney, 2011). In the same way as the Vaughan Williams Memorial Library used their Take 6 project to support their funding bid for the larger Full English project, the Archive could use a smaller LD project to bid for funding. Creating LD for even a small sample could help raise the Archive's online profile and attract users to its other collections.

Digitisation of the collection is important if the Archive is to preserve Fanshawe's valuable recordings, and other areas of the collection. Beyond preservation, a digitised collection is much more accessible and easier for people to use, so the project is worthwhile, and a helpful stepping

block towards LD. The digitisation project comes with challenges: it is time-consuming, expensive, and preserving the files in an appropriate format requires a lot of storage space. But digitisation's still an important process that many cultural heritage institutions are currently tackling. As pointed out by the Vaughan Williams Memorial Library librarian (2019), "if a thing's digitally available, it's only one further step to be able to put them online." This suggests the Fanshawe Archive is doing exactly what it needs to at this stage. And once the Archive is digitised, more options will open up, whether they opt for an LD system or not.

4.3.Collaboration and Support

For big projects, the Fanshawe Archive can get help from outside sources. The Fanshawe Manager is also very willing for people to volunteer with the Archive: "there's plenty of opportunity for people to work here, supporting the digitising – volunteers or partly-paid", (FM, 2019). Ideally a student placement could be arranged with a university or library school. The external libraries interviewed also offer volunteer opportunities for students: the Gerald Coke Handel Library provides their material to universities for supervised students to learn book conservation, and the Piping Centre Library offers student placements. Using students to subsidise library work is mutually beneficial. The students get valuable experience, and the library gets skilled, inexpensive help. PCL (2019) also mentioned that students, being in the midst of learning librarianship, help to keep the institutions they volunteer at up-to-date with library practice, while the institutions can provide references and teach skills.

4.3.1.Collaboration Benefits

The Vaughan Williams Memorial Library stressed the importance of networking, particularly for small institutions: "Just knowing who to go to who might help – that is crucial" (VWM, 2019). The other two small external libraries corroborated, PCL (2019) describing how networking offsets the professional isolation of lone librarians, and GCH (2019) mentioning the informal bartering and mutual support that cultural institutions offer each other. PCL pointed to IAML as particularly important for small music institutions to network; the Fanshawe Manager has actually given talks through IAML to raise awareness about the Archive. As mentioned in Chapter 1, the Fanshawe Archive is actively trying to establish relationships with various institutions, particularly academic.

Both GCH and PCL receive only limited technical support. The Piping Centre computer expert is quite stretched as he maintains the online presence for the whole institution, which includes a shop, hotel and centre website in addition to the library. The Gerald Coke Handel librarian told me that, although they try to get involved in new technologies, limited access to technical support meant that they might be less ambitious than they would like to be (GCH, 2019). Collaboration can help

coordinate online projects that would be infeasible for small institutions to complete alone. The Vaughan Williams Memorial Library's Full English project, which pulled together 19 collections from libraries across the country, is a good example: the VWM is relatively small but was further along the digitisation process than some of the other collections and raised a successful funding bid for their folksong project (VWM, 2019). The VWM ran the project, and hosted the final website, which allowed them to enrich their collections with additional material from the other institutions involved. Meanwhile, the other institutions benefited from their collections being hosted online with the project and some of the other institutions received help from the VWM towards digitising their collections (VWM, 2019).

The external and Fanshawe interviews suggested that smaller libraries can only offer limited support. Both the Gerald Coke Handel and the Vaughan William Memorial libraries indicated that they were able to informally provide advice, talk to people about their projects and the GCH let people see their cataloguing software before making a purchase. VWM (2019) also indicated that they can sometimes direct people to someone in their network who could potentially help them, and they offer a few "very small-scale bursaries". The GCH and the Piping Centre Library have taken on collections from smaller institutions, but generally only collections they were already associated with. Both the GCH and VWM have outsourced some of their digitisation to larger institutions they are affiliated with, which has helped them look towards taking on larger projects, like the VWM's Full English. GCH (2019) told me that they are looking into some digital projects that are still in the early stages. The anonymous national library offers digitisation services for other institutions, but with regards to LD, they upload their data to the Virtual International Authority File (discussed in Chapter 5) but provide no more "bespoke expertise than that" (ANL, 2019).

The Fanshawe Archive currently doesn't have any symbiotic relationships with any small archives or institutions, although the Fanshawe Manager said that they would "quite like" such a thing, perhaps with the School of Oriental and African Studies in London. Villanueva (2016) and Chen (2016) describe the emergence of cultural heritage projects run by indigenous cultures themselves, and so there is scope for the Archive to collaborate with indigenous projects. The Archive already has ties to the National Film and Sound Archive in Australia and the University of Fiji, who have analogue copies of the Pacific Archive. Affiliations with larger institutions that might be interested in Fanshawe's Pacific work could be a good place to start looking for extra funding or support to further the digitisation of the Pacific collection or LD work. If the Archive decides to create a small LD project with some of the digitised Africa collection as discussed previously, they could use it as an example to demonstrate to these organisations where they could go with the Pacific collection and LD. With

these institutions' support, the Archive could have closer access to local indigenous communities who could provide extra perspectives on the materials in the collection.

4.3.2.Collaboration Challenges

The main challenge for the Fanshawe Archive is that collaboration entails giving up some control. It is a very personal collection, and naturally the people involved want to remain true to Fanshawe's vision. But their reluctance to split up the Archive and to have people mess with the tapes limits the number of doors open to them. The Fanshawe Manager mentioned that they turned down the British Library's offer to house the physical collection because it would involve splitting the items by format, and they really want the collection to "stay together in one or two big rooms" (FM, 2019). As discussed in Chapter 1, the physical archive is an artefact in itself, and the potential for it to be housed like a museum exhibition could be interesting for members of the public, so it would be fantastic if the Archive could achieve their goal of keeping the collection together while making it available for members of the public to easily find, view and listen to. But, in the meantime, they may need to turn down potential collaboration projects in order to maintain autonomy, which could hamper their online progress. That said, the Archive is more open to the idea that an online presence could be split, even if the physical archive should stay together, so there may be greater scope for collaboration once the Archive is digitised.

4.3.3.Collaboration and Linked Data

The anonymous national library recently carried out a project where the LD aspects were exclusively outsourced to their partner, who were LD experts. If even a big library outsources LD programming, the Fanshawe Archive would probably need to get outside support in carrying out a large project like LD. The Digitisation section (4.2.2.) suggests the potential for university students to help the Archive out by creating a small LD project as part of a Masters' dissertation or PhD, but collaboration with professional computing engineers could also be extremely useful. Librarians and engineers bring different skills to the LD project. Engineers can provide the technical expertise but talking to the external libraries and Fanshawe Archive demonstrated the knowledge that librarians offer. The Gerald Coke Handel librarian talked about how they have the music and language knowledge to easily decipher unfamiliar music materials. VWM (2019) mentioned that, for the Full English Project, the engineers and librarians brought their different skillsets: computing and cataloguing/indexing, respectively.

While librarians may quite reasonably be reluctant to learn to code LD in a new system, their knowledge of music, the collection, and specialist vocabulary is invaluable. The future of librarianship is likely to be increasingly technological and digital, and the Piping Centre librarian

mentioned that he thinks new librarians will need to have more computing and coding skills than current librarians, adding that librarianship has historically attracted people with humanities degrees (PCL, 2019). Future librarians may find LD coding possible without much technical support, but in the meantime, collaboration is extremely important and potentially mutually beneficial.

4.4.Ethics and World Music Online

The Fanshawe Archive is very aware of the pitfalls for Western collections of non-Western culture. The Fanshawe Manager explained that Fanshawe got permission for all his recordings either through expensive research permits or from local dignitaries and cultural leaders, if appropriate. He also often paid performers for the recording and provided them with copies of the tapes (FM, 2019), taking care to give credit whenever he could to the indigenous people that he recorded. This is pervasive through the journal notes as well as on the Mastertape descriptions. His documentation and accreditation is always extremely detailed, which the interviewee at the anonymous national library mentioned is often not the case for World Music collections, particularly if they are historical (ANL, 2019). Fanshawe also took care to mention performers by name where he could in the documentaries he was involved in creating: for example crediting the Hippo Man, an honorary title for a spiritual leader that Fanshawe recorded and used in his *African Sanctus* composition, by name, Mayinda Orowo, in the documentary that traced his African journey (Fanshawe *et al.*, 2006). The performers' names or the names of their group or region are also always included in subtitles, which interestingly ties in with Roy's (2016) argument that sometimes a piece of art or music can belong to an entire indigenous group, rather than the individual who performed it at one recorded instance. The Fanshawe Manager explained, "when people accuse David of being colonial, taking for the sake of his own means, that's not true. We are preserving for posterity. And this is the 21st century, we are not destroying; David is preserving."

The Archive is also aware that there might be ethical issues that they need to consider if they decide to put the collection online. Both FT(D) and FPT acknowledge that the Archive might contain some sensitive cultural areas, for example, FT(D) mentioned that there are some African recordings where Fanshawe went through some initiation rites, and FPT thought that some of the ceremonies might need careful descriptions that would allow people to study the recording without causing any cultural distress.

An example of a potential ethical consideration for the Fanshawe Archive is the recording of kue kambnibit flutes from the East Sepik region of Papua New Guinea. These sacred flutes are only played during boys' circumcision and initiation ceremony but, after a three-hour discussion with Fanshawe, the village elders agreed to let him record them (Leimbach, 1987). The lengthy

conversation indicates that the elders gave informed consent for the recording to be made, however, the Archive will need to think carefully about sharing that recording online. While the recording is undoubtedly "unusual and rare" (Leimbach, 1987, p.26), and therefore valuable, the details of Fanshawe's agreement with the village elders would have to be examined to ensure they had given permission for such widespread propagation of this sacred ritual instrument. Moreover, the risk of a recording being pirated by users is greater when the recording is online, and should be weighed before spreading the recording beyond the Archive's sole control. The Archive may or may not discover that the appropriate permissions were granted by the village elders at the time, but either way, the ethical consideration must be undergone, and it cannot be assumed that permission to record is the same as permission to post online. As ANL (2019) put it, "I think you have to be really careful about making assumptions about opening it up to the world when perhaps that's not appropriate to the culture that it came from." The Fanshawe Manager (2019) told me that recordings without a known performer are legally in the public domain, but the ethicality of sharing the recording may still be questionable. Nevertheless, from our conversations I think the Archive realises this and will be careful to present their materials respectfully, if they decide it is appropriate at all.

4.5.Benefits, Challenges and Reception of Linked Data for Small Archives

Currently, the Fanshawe Archive is not in a position to focus on LD, but they are still making very important steps towards making their collection more accessible online. In future, they may or may not adopt LD. While LD would have benefits, there are also challenges that are compounded by being small and having limited resources.

4.5.1.Benefits of Linked Data

Interviews with the Fanshawe Archive and external libraries highlighted benefits that LD could bring their online catalogues and collections. The Fanshawe Archive would be interested in LD if it would help their users to relate items within the Fanshawe collection to similar resources either within their own catalogue or externally (FM, 2019; FT(D), 2019; FPT, 2019). The Fanshawe Manager recognised that the Archive needs a greater level of automation for finding the items in the collection; they can't rely on their current, somewhat time-consuming "manual research" methods of searching through organised boxes to find a reference to determine which journal entry relates to the sound recording, etc. "You should be able to see in a main glance where all the links, where all the branches are" (FM, 2019). FM mentioned that she doesn't know enough about LD to make a judgement on its benefits, but I feel that LD could help the Archive interconnect the resources within the collection itself and because FM indicates that she wants the Archive itself to be more coherent,

the first example of Chapter 5 examines how they can use LD to join up related materials in their archive.

The part-time member of staff, who had researched LD in preparation for my site visit, was interested in the sharing aspects, but also the increased browsability that LD offers, since it "makes it so much easier for people to see where relationships are that they might not have previously considered" (FPT, 2019). The anonymous national library interviewee, who had participated in a project involving LD, similarly emphasised that LD allows users to see the different relationships between disparate items, adding that it makes it easier for users to search across different organisations for a topic, and across various sources from separate organisations, which is "something that would be quite hard to search for using the normal catalogue interface, normal internet and normal data structures" (ANL, 2019).

The Piping Centre Librarian (2019) also mentioned the greater resource discoverability offered by LD, both making their current users aware of what else is out there and making their collection more discoverable online. Greater traffic to the Piping Centre Library catalogue would raise the profile of the Piping Centre itself, which would bring revenue to the centre. The Fanshawe Collection similarly has a separate business arm, which is not part of the trust, but if LD raised the Archive's online profile, this might generate some financial return for the Fanshawe Business.

The Vaughan Williams Memorial Library recently completed the Full English project, which placed English folk music, folk dance and folk art in their wider context. Although they didn't use LD, the project's aim was "to make stuff more available by presenting it online and interpreting it" (VWM, 2019). This is exactly what LD facilitates, so it is significant that even outside the Linked Datasphere, libraries are finding ways to make connections between resources. The role of libraries and other cultural heritage curators is a vital part of the process since, as VWM mentioned, it's not enough to just present things online, the interpretation of the material is equally important. This relates back to the discussion on collaboration, where both librarians and technical experts bring valuable skills to online projects, including ones that use LD.

The anonymous national library also talked about the benefits of LD for smaller libraries in particular. LD would put the smaller institutions' collections in context and raise their profile by showing how they fit into a wider picture. ANL (2019) also mentioned that LD has the potential to challenge the current dominance of larger institutions by "showing how smaller institutions also have important collections with important stories to tell."

4.5.2.Challenges of Linked Data

LD is a big project and the interviews identified three main challenges that a small institution like the Fanshawe Archive would face in implementing an LD system: lack of time, lack of money and lack of technical expertise.

There are five people involved in the Fanshawe Archive, and only the Fanshawe Manager works fulltime. FT(D) (2019) mentioned that he finds the amount of work involved in cataloguing even without the added complication of LD "kind of mindboggling", and he questioned where they would find the time to input LD programming. FPT (2019) echoed this concern, stating that, although "it would be great" to take the time out to do something with LD, limited time and money would be "the main stumbling block." Even the anonymous national library's LD project, that involved input from many more people than the Fanshawe Archive has access to, was difficult to complete within one year. In fact, they could not implement the final project on their own library website in that time, so it is currently hosted externally (ANL, 2019). Similarly, even with the support of an institution like the Piping Centre behind it, the Piping Centre Librarian told me he is probably not in a position to take on big overhaul projects like LD. Lack of time and staff resources make it difficult for smaller organisations to consider LD as a viable option, and PLC (2019) referred to LD as "a bit of a step change."

On top of the time it takes to input LD coding, the learning process for this also takes time and can be very daunting. PCL (2019) talked about the challenges for librarians, particularly lone-working librarians, to keep up with the tech world. The tech world is moving so fast that, while institutions like the Fanshawe Archive are still getting their resources into a digital format, some researchers and libraries are jumping ahead to look at ways to move to the next step beyond LD, discussed in more detail in Chapter 5. These smaller institutions are continually forced to play catch-up, with fewer resources at their disposal.

The Fanshawe Archive would probably need to apply for external funding if they were to try to move on to LD. FPT (2019) told me that an LD project would probably require a specialist expert's help, but she mentioned that once the Archive is further along with the digitisation project, they would be more appealing to potential LD sponsors – "we could be pioneers in it!" This returns to my earlier argument that the Archive's biggest asset is its collection and a digitised collection would be more attractive for external funders. The anonymous national library's project was expensive, but it included extra factors outside LD and much of the cost went on digitising resources, as well as staffing (ANL, 2019). A less ambitious project would cost less, particularly if the digitisation had already been done. The Fanshawe Archive's current involvement with the Save our Sounds project

will complete their Africa collection. As discussed above, a small LD project that focused on the African Journey, for example, would cost less, as well as be more feasible and a good demonstration pilot when bidding for funding.

The Fanshawe Archive is a private collection, and therefore, like the Gerald Coke Handel Library, is ineligible for some funding that is reserved for higher education institutions (GCH, 2019). Happily, though, cross-sectoral funding is becoming more widespread (GCH, 2019). The Vaughan Williams Memorial Library's Full English Project was sponsored by the Heritage Lottery Fund (VWM, 2019), and the anonymous national library's project got support from the Arts and Humanities Research Council as well as partnering experts from a university (ANL, 2019). While the Fanshawe Archive could apply for funding bids they are eligible for, it might be more feasible for them to keep an eye out and advertise that they were interested in being involved in relevant LD projects led by larger institutions rather than trying to tackle such a project alone.

The anonymous national library identified that there are no "off-the-peg" cataloguing systems that support LD (ANL, 2019). Such a system, if created, would help lessen all three of the main concerns surrounding LD. A user-friendly interface for cataloguers would mean that the technical challenge of learning complicated coding would be alleviated; a user-friendly cataloguing structure that guided the librarian in RDF would save time and help keep things standardised, and if the program were user-friendly and designed around librarians' existing skillsets (as opposed to the skills of computer programmers) the librarians could input the data themselves without having to spend money hiring LD experts. It would make LD adoption more feasible for small institutions like the Fanshawe Archive. However, until the creation of such an off-the-peg LD system, small archives must rely on collaboration with libraries and other cultural heritage groups to keep up with the fast-paced technological advances of the wider world.

4.5.3.Overall Reaction to Linked Data

The overall reaction to LD from everyone I interviewed was generally positive in theory, but there was trepidation over how feasible it would be and insecurity over technical skills. The Gerald Coke Handel librarian mentioned that they had learned some coding when uploading their manuscripts to RISM UK, and because they were given proper training, it was something they were technically capable of. After being given a description of LD, they said that they would "definitely" be interested "if you could explain it", and, while citing that they only have two staff and therefore limited time, added later "if someone showed us how to do it, we'll do it."

The Vaughan Williams Memorial Librarian hadn't heard of LD, but although they didn't use LD in their recent project, the Full English, that project does link library resources to outside sources. For

example, some of the Broadside ballads link to online images of the Broadside hosted by the Bodleian Library at Oxford. They also link internally to relevant images and, more recently, sound files – but again, not using LD.

Although unsure whether an LD project would be feasible for the Piping Centre Library any time soon, PCL (2019) mentioned that he thinks that national libraries really ought to be using LD since, "if they're not doing it then there's something far wrong with the way that libraries are being led." As discussed further in Chapter 5, national libraries are the trend setters for librarianship, and the anonymous national library I interviewed is using LD, with support from external experts.

A recurring theme across the interviews was confusion and a limited understanding about what LD was. Some interviewees indicated that they knew nothing about LD, while others thought they might have known about it under a different name. Some initially thought they were already using LD but once they found out more about it, realised we were talking about different things. Those who were unfamiliar with the term assumed that it was to do with linking resources to others using hyperlinks or shared online catalogues, which indicates that the term 'Linked Data' is perhaps unhelpful when trying to attract libraries to it. Perhaps the boat has sailed on renaming it (and I don't have an alternate suggestion), but advocates should be aware that the term can be confusing to those who are unfamiliar with it. Hopefully, as LD becomes more widespread, any confusion about the name will become less of an obstacle.

It is encouraging that all the institutions interviewed, including the Fanshawe Archive, are keen for their library to be more interconnected. Some are already making steps towards this, but they are limited by a lack of money, time and staff. The technical aspects of LD are viewed quite apprehensively, and the RDF framework is generally unfamiliar, particularly regarding XML coding. The national libraries are leading the way in demonstrating how LD can be used in libraries, although there are still challenges in making smaller institutions aware of how LD would benefit them. LD may seem like an insurmountable task for small institutions like the Fanshawe Archive, but the interviews with the Archive and external libraries gave a sense of the smaller groups of talented individuals pulling their expertise together to make their collections as useable and accessible for their users as possible.

The Fanshawe Archive's current digitising and cataloguing projects are huge, and an LD project would also be very big, but this chapter has demonstrated the value of LD, while acknowledging the practical challenges and ethical considerations for a small World Music archive. I have made some suggestions about what the Archive could implement now in their current projects to make a future LD project less daunting.

CHAPTER 5 – ANALYSIS 2: LINKED DATA AND THE FANSHAWE ARCHIVE

This chapter takes examples from the collection to examine how the Fanshawe Archive might start creating an LD catalogue. The following Maps and Graphs use my understanding of the resources to demonstrate ways the Archive might describe the sound recordings and the process to use LD with this collection. A full project would be much larger, involving technical experts and the people at the Archive, who have a deep knowledge of the materials and context.

Pennington (2019, p.157), discussing the creation of a dementia ontology, advocates centring the ontology on people living with dementia, since it "reminds everyone to keep the PWLD as the focus." Similarly, FT(L) (2019) believes, "the heart of the collection is the recordings and everything else is wrapped around that." FM (2019) likewise talked about structuring the Archive catalogue: "where all the branches are the components for this one recording, with the recording as a focus."

The two examples below, therefore, focus on the sound recordings and how they can be linked. Example 1 looks at FAC 44, the Bunyoro Madinda, and how the Archive can use LD to establish connections between items within the collection, sometimes using external URIs as concepts that resources have in common. Example 2 (FPC 1389, the Wagi Brothers Bamboo Band) demonstrates how the Archive can link outwards to external sources, to enrich users' experience of the catalogue and place the recording within its wider context.

5.1.Example 1: Bunyoro Madinda (FAC 44)

This example demonstrates how LD can be used to show relationships between the collection's items. It also demonstrates the theory of creating an LD catalogue entry. The big picture is first presented in a Concept Map (Fig.5.3.), written in human language. The Map shows the kinds of links, both internal and external, that could exist in an LD system to make the Fanshawe collection more coherent. Fig.5.4. then directly translates the previous Concept Map into RDF language using ontologies and URIs; I'm calling the computer language version a Graph. A computer still couldn't read it, but the process allows a direct comparison between the human-centric Concept Map and the computer-focused Graph. Up to this point we'll have been looking at the wider picture, so Fig.5.5. narrows the focus to just the direct branches from the central Madinda Catalogue Entry nucleus. These branches represent the fields that would be used in the final catalogue entry for the Bunyoro Madinda, although the full entry would have more fields/branches. Fig.5.6. and Fig.5.7. then demonstrate the layout for some of the final code, coding the semantic relationships into XML so the computer can understand them.

5.1.1.Creating a Concept Map

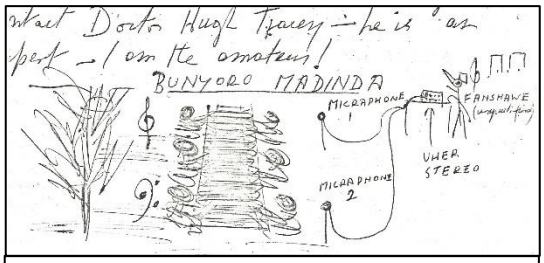

Fig.5.1. Fanshawe's Madinda Sketch
© Drawing David Fanshawe

Fanshawe's notes describe the Bunyoro Madinda as a line of boards laid across banana trunks with sticks poking through to keep the boards in place. It's a bit like a xylophone played by six performers. The notes, in this case, don't come from a journal; FM found Fanshawe's description among photocopied letters sent to the Vaughan Williams Trust. This demonstrated to me how valuable FM's knowledge of the Archive is when pulling together all the Archive's resources on any individual recording. Due to her familiarity with the Archive and Fanshawe's travel history, when we couldn't find anything about the Madinda in the journals, FM was sure there was something somewhere, and sure enough, we found it among the letters. The letter we found is very interesting from a musical, a historical and a sound engineer perspective. It describes the instrument, gives more information about the players, and includes a drawing by Fanshawe of the setup of the instrument, performers and microphones (Fig.5.1.).

In terms of sound material on the Madinda, the Archive has the initial recording on analogue tape and they will soon have the digitised version. There are 5 slide photos (already digitised) of the Madinda, including one of Fanshawe recording the instrument, and FM pointed me towards related photographs of the drum that accompanies the main instrument on the recording, a photograph of Lake Kyoga, where the recording took place, and photographs of similar instruments that Fanshawe recorded such as the Marimba Xylophone from Tanzania and the Pong Lang Xylophone from Thailand. Additionally, the Archive has scans of the Masterbox, which includes extra information such as the location (Lake Kyoga) and the fact that the performers were fishermen.

Therefore, even within the Archive itself, many items could be linked up just for the Bunyoro Madinda. Although the literature on LD concentrates on the external links that would connect collections together, the nature of the Archive, with so much related material, lends itself to internal linking as well.

To take an example outside the Madinda, the Fangufangu nose flute recordings are much more interesting when linked to the photographs of the Honourable Ve'ehala playing the instrument (Fig.5.2.) due to its unusual way of being blown through the nose. While pretty – the recordings just sound like a soft flute – at least in a Western context, the main interest comes from the unique playing style. So, creating a link between the two would be useful for browsers of the Fanshawe

Archive catalogue and also those who were looking for nose flutes or similar relevant material online.

Fig.5.2. The Honourable Ve'hala playing Fangufangu nose flute.
© Image David Fanshawe

Fig.5.3. (overleaf) shows the start of a Concept Map using information about the recording from the letter, the original Masterbox, and a Google search. A finished Map would contain more branches – the Map here only links to one photograph of the Madinda, for example – but this serves to demonstrate some of the links that the Archive might make, and also shows how items within the collection (pink boxes) can be linked directly or through external concepts (blue boxes).

A more common name for the Madinda is actually Amadinda (Kubik, 2010; Bartolome, Mukuna and Oehrle, 2010), and although I first discovered this on Wikipedia, I was able to corroborate the information in several books, one of which Wikipedia references. That said, I realise in practice it may not be feasible for the Archive to verify every piece of information either that Fanshawe wrote or that they discover in external sources. An advantage of the collaborative nature of LD is that external experts can contribute to linking information (Spiteri, 2019), so even where no-one within the Fanshawe Archive had recognised the link between the two, it is likely that if Web 3.0 were to become more pervasive, someone would eventually notice and make the connection.

LD can also help overcome language barriers. Since Fanshawe visited so many places, it would be impossible for the people at the Archive to know all the local languages on the recordings. FPT (2019) mentioned that, for example, transcribing the journals might not be possible in all cases. This stands in contrast to a Western collection like the Gerald Handel Coke Library where, between them, the librarians have the European language skills necessary to translate the material they encounter (GHC, 2019). The collaborative nature of LD means that indigenous communities could aid with the resource description and translation of materials.

The above example includes a hypothetical individual catalogue entry for the letter that Fanshawe wrote about the Madinda. However, the Archive might decide that creating individual entries and URIs for each letter is too onerous for any gain it might offer users. The external interviews with VWM and ANL showed that wider and narrower cataloguing approaches are both valid, depending on the goals of the institution and needs of users. As mentioned in Chapter 4, the Vaughan Williams Memorial Library chooses to catalogue its online collection down to the individual item, where the larger national library elected to group related ephemera into single cataloguing entries. Like with

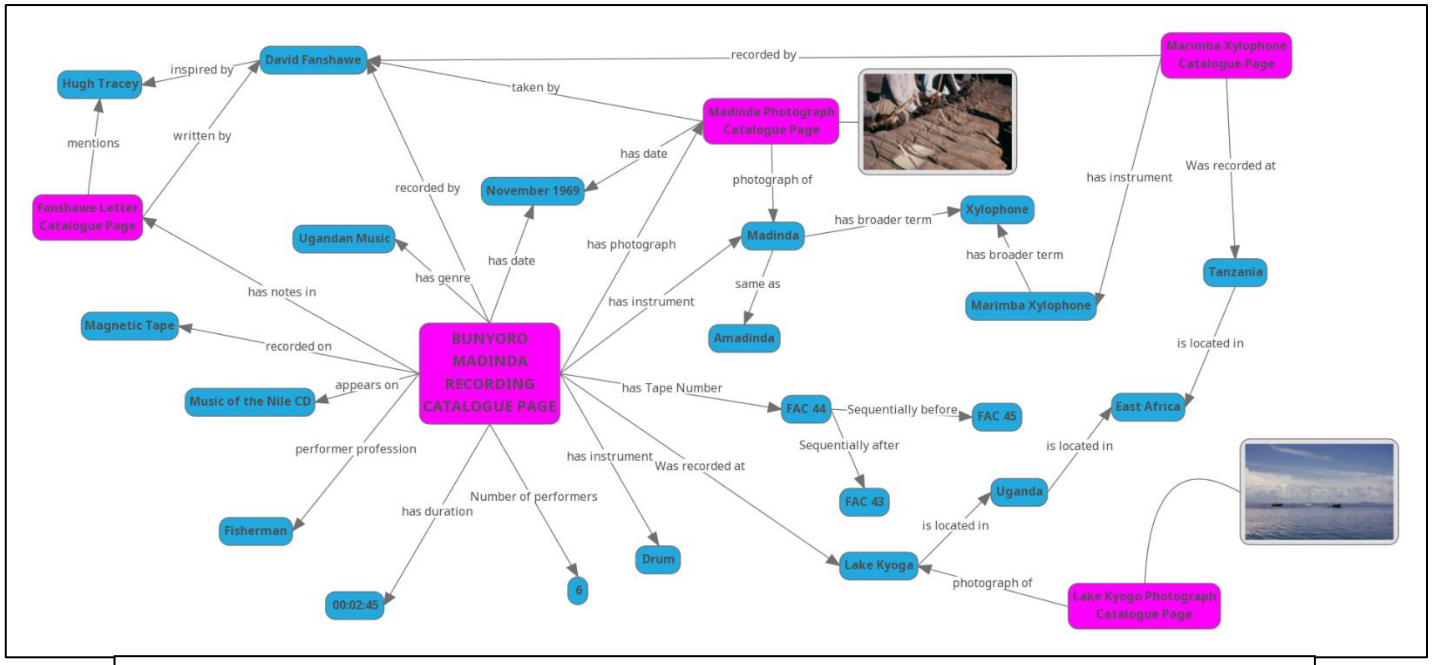

Fig.5.3. Example Concept Map for the Madinda Recording (FAC 44)
© Images copyright David Fanshawe

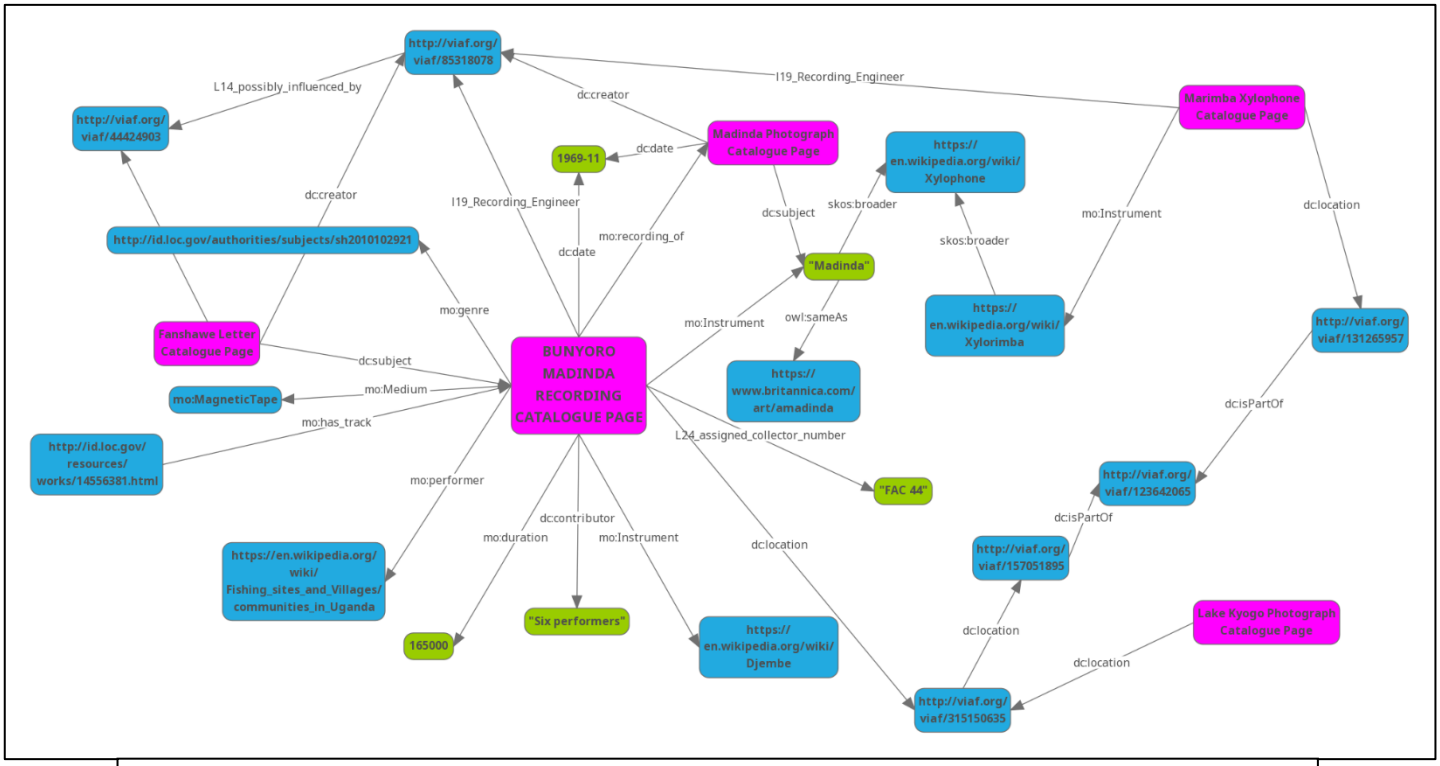

Fig.5.4. Example Graph for the Madinda Recording (FAC 44)

Fig.5.3. and Fig.5.4. have been laid out together in order to make it easier to follow the transition from human-friendly language to computer language, which uses ontologies and web addresses. The Figures are laid out to correspond as closely as possible to each other, and larger (easier-to-read) examples are included as Appendices.

the journals, it might be most helpful and immersive for the Fanshawe Archive to create a single entry collecting all the letters to the Vaughan Williams Trust, with deep indexing to point users to the appropriate pages.

5.1.2.Creating a Linked Data Graph

The Concept Map in Fig.5.3. is the first step, but a Graph (Fig.5.4.) only seems complicated because it uses ontologies and URIs, rather than words that humans can better understand. This one corresponds directly with the above Map, which is the kind of thing the Fanshawe Archive could create right now, with no coding expertise, and the deeper familiarity that FM and the others have with the collection would result in considerably more branches and connections.

The main consideration for the Fanshawe Archive, after deciding the links they want to create, is to select which ontologies to use. The Archive is sceptical that existing resources would fit their varied and unique resource description needs, and are therefore creating a thesaurus (FM, 2019). Ontologies and thesauri are similar, although ontologies structure the knowledge to be machine-readable, using "strict semantic relationships" (Harpring, 2010, p.25), and finding an ontology that suits a World Music collection like the Fanshawe Archive's has the same complications. Weissenberger (2017, p.2), speaking with regards to traditional Irish music, states that, "Among the few music ontologies developed, such as the Music Ontology, none adequately express orally-based traditions like Irish traditional music and dance…The majority…are based upon the norms of Western Classical and Western Popular music."

My research has not unearthed anything to contradict the apparent lack of designated World Music ontologies, either that Weissenberger missed or that have been created in the interim. Weissenberger (2017) cites three music datasets that are not grounded in Western Classical music (the Linked Jazz project, the Europeana Sounds project and ethnoArc) but each would be unsuitable for the Fanshawe Archive's collection. Linked Jazz is an ontology of Jazz music; Europeana, as suggested by the name, focuses on European cultural heritage; ethnoArc is a primarily euro-centric folk archive but, crucially, is not built for LD.

If a World Music ontology does exist that we have both missed, then it is quite obscure. While this might pose fewer problems for a larger institution with more trend-setting potential, for a small archive, implementing a perfect dataset that barely anyone uses could undermine their attempts to join the Linked Open Data (LOD) Cloud. An institution like the Carnegie Hall Archive has enough weight to justify their project to create "canonical" URIs for every performance at the Carnegie Hall Theatre (Hudson, 2015), but a comparatively less well-known institution like the Fanshawe Archive might struggle if the Linked Dataset they chose was practically unfindable.

Therefore, I recommend that the Fanshawe Archive primarily uses the Music Ontology since it has a variety of relevant fields and is already widely used by other music-related LD projects. Kelly (2015) asserts that, "the proliferation of the Music Ontology through these other projects contributes to the goal of creating a semantic web" (p.4). For the Fanshawe Archive to have the best chance to align with the larger institutions and the semantic web, it therefore makes sense for them to use a well-known pre-existing dataset.

This is a recommendation with reservations, however, and there are a few challenges that the Music Ontology creates that might be alleviated by the creation of a more overarching and encompassing ontology. The Music Ontology is designed to work alongside other vocabularies, particularly FOAF (Friend of a Friend) (Raimond *et al*., 2013), which is useful because it means that the URIs are more standardised, but unhelpful for coders since they have to go to several different places to find suitable vocabularies. Zhao and Ichise (2014) argue that the vast number of different ontologies, datasets and RDF triples already in the LOD Cloud makes it extremely time-consuming for programmers to manually sort through all the information to find and use ontologies properly. The extra time required looking for vocabularies and ontologies exponentially affects smaller institutions like the Fanshawe Archive.

Although designed to be combined with other ontologies, the Music Ontology does not fit ideally with orally transmitted music (Weissenberger, 2018). To remedy this, the Irish Traditional Music Archive has created an ontology, named LITMUS, that is more suited to the oral tradition shared by many non-Western classical music styles (Weissenberger, 2018). At least until the creation of a designated, widely-used World Music ontology, I suggest that, in cases where the Music Ontology falls short due to style differences between classical and folk music, the Fanshawe Archive turn to LITMUS to look for appropriate terms. Additionally, a Performed Music Ontology has recently been created (Futornick, 2019), and since the Fanshawe collection contains many Mastertapes of live music performances, this might also fill some gaps – particularly for Fanshawe's Pacific music festival recordings. Beyond that, FT(L)'s cataloguing experience will help with the selection of further appropriate, more generic, ontologies that suit the collection. Alongside FOAF, the Music Ontology recommends Dublin Core (Raimond *et al.*, 2013), which is already used by digital libraries, such as the Taiwan Digital Archives, to describe indigenous artefacts (Chen, 2016), although not in an RDF structure.

Although the Fanshawe Archive have indicated their willingness to create a thesaurus for their catalogue (FM, 2019; FT(L), 2019), the creation of a World Music ontology (ideally one that is compatible with the Music Ontology and other larger datasets) is a huge task that would be

impossible for a single small archive like the Fanshawe Archive. It seems more feasible for the burden of creating a World Music ontology to be taken on by larger institutions such as a national library or a music librarian association like IAML or MLA. If a larger group were to attempt to create one, looking at the LITMUS datasets would be a good starting point. Although designed for Irish traditional music, Weissenberger (2017, p.1) mentions that the LITMUS ontology could "serve as a reference point for other Linked Data projects involving orally-based music traditions." She later recommends that "using practitioners' own language will benefit the application of the ontology within traditional music collections in Ireland, as well as when applied to other European and non-European music collections" (Weissenberger, 2018).

5.1.3.Coding Linked Data

The previous two sections demonstrate the wider picture, showing how the Madinda example relates to other items in the Fanshawe Collection and then directly translating it into a Graph, using URIs and ontologies. In practice, each item's catalogue record (URI) requires its own RDF coding, in the same way that each catalogue record has a MARC counterpart. Fig.5.5. cuts the Graph in Fig.5.4. down to show the direct branches that could be fields within the Madinda catalogue.

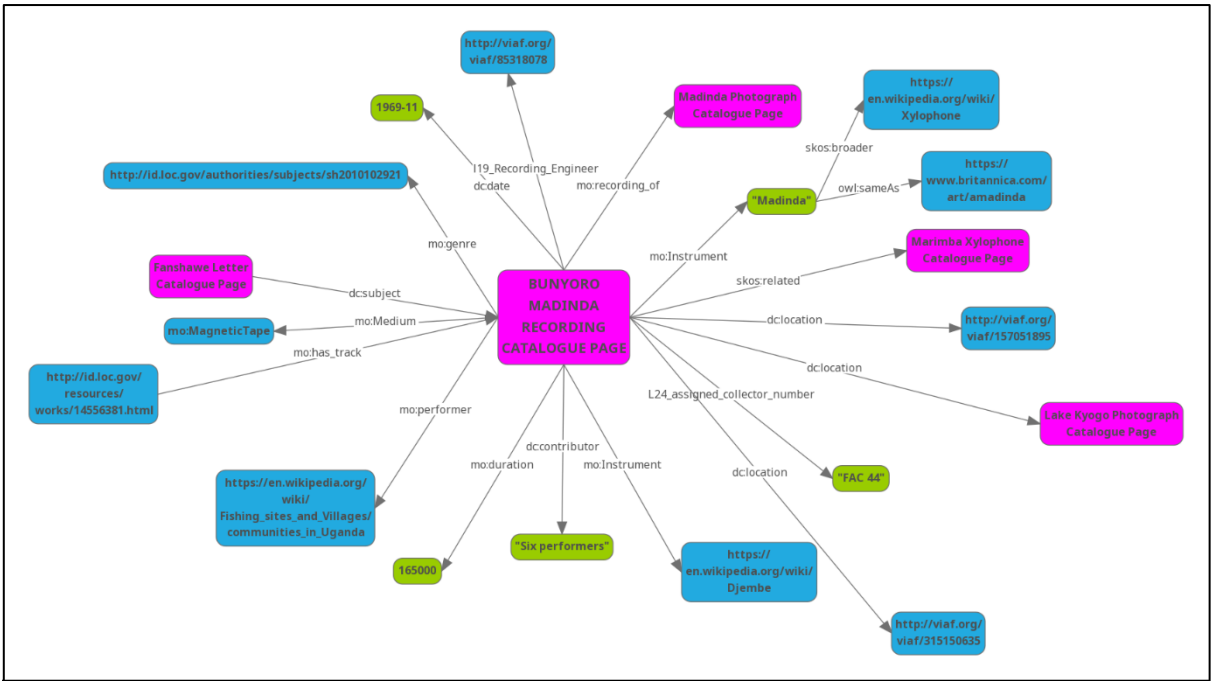

Fig.5.5. Example cut-down Graph for the Madinda Recording (FAC 44)

With this reduced Graph, all that remains is to write the code. The following coding example uses RDF and starts with prefixes to define the ontologies used (Fig.5.6.). Since RDF is designed to "mix and match terms from multiple vocabularies", which the Music Ontology relies on heavily (Raimond *et al*., 2013), it would make sense to use RDF with the ontology even if RDF weren't the structure that LD centres on. The final coding (Fig.5.7.) follows the same RDF format throughout: Subject → Predicate → Object.

```
@prefix mo: <http://purl.org/ontology/mo/> .
@prefix dc: <http://purl.org/dc/elements/1.1/> .
@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .
@prefix tl: <http://purl.org/NET/c4dm/timeline.owl#> .
@prefix event: <http://purl.org/NET/c4dm/event.owl#> .
@prefix foaf: <http://xmlns.com/foaf/0.1/> .
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
@prefix viaf: < http://viaf.org/viaf/> .
```

Fig.5.6. Coded prefixes for ontology namespaces

```
<http://FanshaweBundoroCatalogueURI> <L24_assigned_collector_number> <"FAC44"> .
<http://FanshaweBundoroCatalogueURI> <l19_Recording_Engineer > <http://viaf.org/viaf/85318078> .
<http://FanshaweBundoroCatalogueURI> <mo:genre> <http://id.loc.gov/authorities/subjects/sh2010102921> .
<http://FanshaweBundoroCatalogueURI> <mo:Medium> <mo:MagneticTape> .
<http://FanshaweBundoroCatalogueURI> <mo:duration > <"165000"> .
<http://FanshaweBundoroCatalogueURI> <mo:Instrument> <https://en.wikipedia.org.wiki/Djembe> .
<http://FanshaweBundoroCatalogueURI> <mo:Instrument> <https://ww.britanica.com/art/amadinda> .
<http://FanshaweBundoroCatalogueURI> <mo:recording_of> <http:FanshaweMadindaPhotographCatalogueURI> .
<http://FanshaweBundoroCatalogueURI> <dc:date> <"1969-11"> .
<http://FanshaweBundoroCatalogueURI> <dc:contributor> <"Six Performers"> .
<http://FanshaweBundoroCatalogueURI> <dc:location> <http:FanshaweLakeKyogoPhotographCatalogueURI> .
<http://FanshaweBundoroCatalogueURI> <dc:location> <http://viaf.org/viaf/1570522895 > .
<http://FanshaweBundoroCatalogueURI> <skos:related> <http:FanshaweMarimbaXylophoneCatalogueURI> .
```

Fig.5.7. Final piece of code, written in RDF and colour coded to correspond to Subject → Predicate → Object

After the harder work of selecting ontologies, finding vocabulary and deciding what links to make, the final coding just follows the same computer grammar throughout, so isn't too challenging. The most time-consuming part is finding suitable ontologies, and a World Music Ontology would make finding appropriate vocabularies much easier.

5.2.Example 2: The Wagi Brothers Bamboo Band (FPC 1389)

Example 1 showed how items within the Fanshawe Archive's collection could relate to each other and demonstrated the kinds of links that would help connect the items to each other. If the Archive used RDF to only create links within its own silo, this would still push it to four stars on Berners-Lee's (2009) 5-star Linked Open Data development scheme. Helping the computer to understand the collection's items in relation to each other would already be a big step in the right direction.

The final step, linking outwards to other useful data, is really the main goal of LD – putting the collection in its wider context and making it more retrievable on the web. The World Wide Web Consortium recommends that dataset publishers "select data that is uniquely collected or created by

[their] organisation. Ideally, this information when combined with other open data provides greater value" (W3C, 2014). Although some of Fanshawe's recordings have been released on CDs, the Fanshawe Archive is the hub of this unique collection. Below are some examples of how the Archive might link outwards to join the LOD Cloud.

David Fanshawe already has a Virtual International Authority File (VIAF). VIAF uses LD (OCLC Developer Network Team, 2019), so the file will already have been used as a URI by other institutions and so can be used as a URI by the Fanshawe Archive as well. This is helpful for the Archive, because everything in its collection will have some connection to David Fanshawe. Even if an RDF connection to Fanshawe's authority file is the only external RDF link they make, everything in the collection will be connected to the LOD Cloud. Although I'm not suggesting they take this route, it demonstrates how easy it would be for them to join up with the LOD Cloud from the outset. What's more, WorldCat uses VIAF, meaning that the Fanshawe Archive's collection could join up with the world's largest union catalogue through these links. Coyle (2010) argues that union catalogues like WorldCat remain siloed from the rest of the web. However, the fact that WorldCat uses VIAF signals that LD is becoming more widespread and it is likely that WorldCat and other union catalogues will adopt LD approaches in future, if they haven't already.

Fanshawe's VIAF is an area where it will be easy to create an external LD link, but what about some of the more obscure fields? For example, the Archive's recording of Gilo Stones from the Solomon Islands seems to be the only recording of this particular instrument. The stones are played by being rhythmically hit by hollow sticks, so could perhaps be linked to the URI for percussion instruments, but there seems to be no existing URI for the instrument 'Gilo Stones'. In fact, the only online mention outside the Fanshawe recording (which appeared on a few CDs), was amidst a glossary of instruments from a folk instrument shop in Arizona (The Folk Shop, 2019), and this didn't use LD. In this particular case, the recording could be linked to stores or libraries that carry CDs that it appears on, but this only applies to a small percentage of the recordings in the collection. It is likely that some items that won't even have that connection, which demonstrates the equal importance of links between items in the collection.

Some recordings appear in DVDs or Radio Broadcasts held by the Archive. For example, the Bwala dancers recording features on a radio broadcast (Fanshawe, 1973) and in Fanshawe's composition *African Sanctus*. Likewise, a recording of a rock bashed against a resonant stone on the mouth of a volcano appears in a TV programme about Fanshawe's Pacific journey. This resonant stone is called a bell stone, and the instrument exists across islands in the Pacific and Hawaii (Fanshawe, 1987), so as well as linking to the programme, links could also be created to other bell stone recordings or sites.

It was sometimes surprising how much related material a recording could be linked to. I was fascinated by the Wagi Brothers Bamboo Band because of its odd instruments: huge, hollow bamboo tubes lined up and struck with plastic flip-flops. It turns out, however, that bamboo bands are an established style in the Pacific. The Solomon Islands, having only been released from British sovereignty in the 1960s, has a unique brand of Western-inspired, but still distinctly Pacific, music (Ellingham, Duane and McConnachie (eds.), 2000). Fanshawe's recording is probably one of the first to capture this fantastic style in its early years, but since then, there have been other bamboo bands, such as the Kalibobo Bamboo Band in the 1980s and the current Pynolasa Bamboo Band. The latter appears on WorldCat and has some LD associated with it. Discovering this only required a brief Google search, so shouldn't be too difficult or time-consuming for a cataloguer. Crucially, though, a reverse search into bamboo bands returned nothing about the Wagi Brothers Bamboo Band, although the Pynolasa Bamboo Band was mentioned – suggesting that LD could increase the Wagi Brothers' recording's accessibility as it has for Pynolasa.

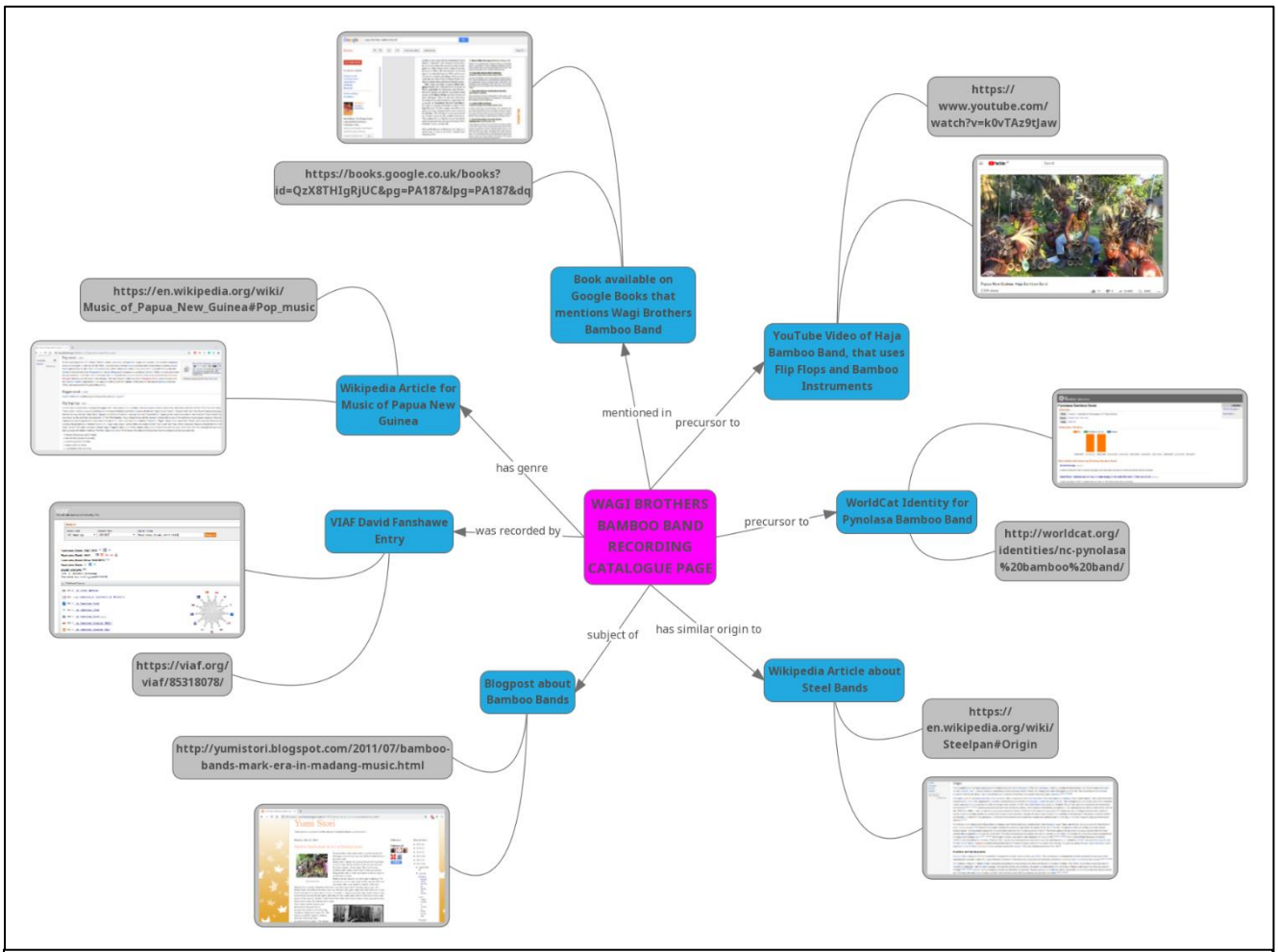

Fig.5.8. Example external links for the Wagi Brothers Recording (FPC 1389)

The external links displayed in the Concept Map above (Fig.5.8.) include a link to an online World Music book that contains the information about the Solomon Islands' unique blend of Pacific and Western music and to the other two bamboo bands mentioned above. The book also describes similarities between the bamboo bands and steel drum bands of the Caribbean, where instead of plastic flip flops, the Western debris that locals repurposed into instruments were oil drums (Ellingham, Duane and McConnachie (eds.), 2000). Links could also be drawn between the bamboo bands of Polynesia and Asian bamboo bands of China, for example, which are possibly more familiar to a Western audience.

Some of the music in the Fanshawe Archive itself will be more familiar to a Western audience than other recordings, and links from these to external sources (and vice versa) could help as a jumping off point to make the Archive's collection more relatable to Western users. In the Pacific Islands, for instance, Fanshawe often travelled between churches to record congregations (Fanshawe, 1987), and some of the music will be recognisable to people familiar with Christianity. For example, at the Archive I listened to a recording of a Methodist choir in Nabua (Philippines) singing the popular hymn *Whispering Hope* in their native language with accompanying drumming. This recording could perhaps be linked out to that hymn, or to further information about Methodism in the Pacific Islands.

In Fiji, Fanshawe also recorded some Methodist hymns of the *sere ciri* genre, i.e. composed by anonymous native Fijians and not found in the official hymn book (Fanshawe, 2002). The attempts to insert *sere ciri* choruses into hymn books in the 1980s were cut short when the Methodist church stopped printing books that included the choruses (Qiolevu, 2015). There have, however, been recent movements in Fiji to revise the hymn book to include compositions by native Fijians (Qiolevu, 2015), so the recordings might have some valuable contemporary cultural relevance for a wider Fijian audience. A *sere ciri* URI would therefore be useful for describing the Fiji recordings in LD, since it might connect Fanshawe's recording to the contemporary Fijian movements. Sadly, the rarity of examples of that genre, particularly online and especially in the West, means that without a dedicated World Music ontology, it's unlikely that such a URI exists. The recordings could still be linked up with each other, the Philippine recording and/or the wider topic of Methodism in the Pacific Islands. This example, and the Gilo Stones example demonstrate that, for LD to properly accommodate a World Music collection like the Fanshawe Archive, a World Music Ontology should be developed, or at least another music-based ontology extended to be more World Music inclusive. The Fanshawe Archive, with its wealth of unique recordings and detailed information on rare instruments and styles would be a valuable resource to anyone attempting to create such an ontology. As discussed in Chapter 4, a World Music Ontology project could perhaps gain access to

the Archive's knowledge and collection in return for offering expertise in creating LD for the Archive's catalogue.

As mentioned in Chapter 1, the Archive has quite a small userbase, comprising mostly of commercial users or students searching for a particular World Music specialisation. FM (2019) mentioned that one of the topics students might look for is the gospel hymnary chanting of Polynesia. A student of the subject might find technical resource-description terms, like *sere ciri*, helpful since this accommodates the narrow, specialist search that they might conduct, and Fijian natives might also find the technical description useful if the term is more common or culturally relevant in the Pacific.

The Archive are keen to extend their userbase, and reaching a wider audience is something LD is designed for. The VWM librarian (2019) told me that they receive more online visitors than visitors in person to the library – from complete beginners to experts. This shows that for their online presence, the Fanshawe Archive needs to consider how they will accommodate non-specialist browsers, who might be turned off by excessive technical language. The anonymous national library had style guidelines for presenting information. While their material is intellectually rigorous, they are encouraged to avoid jargon and explain technical terms "to keep everything accessible as possible" (ANL, 2019). Currently, where possible, the Fanshawe Archive uses terms based on Fanshawe's journals and notes, which themselves drew on conversations with indigenous people (FPT, 2019), but they also want to make their descriptions accessible to a wider audience. FT(L)'s library experience is valuable here because she already understands the theory behind resource description. FT(L) (2019) commented, "there's got to be more information for people to track down these different items," going on to explain that they are looking into identifying terms that would be most useful and consistent.

FT(L) mentioned that with resource description and cataloguing you must stop somewhere, and with LD (like most schemas) there is the trap of wasting time rabbit-holed in one knowledge branch. The Archive has 40,000 items and five members, so they are limited in how extensive their linking for each item can be. The Archive doesn't need to create all their external links immediately. As mentioned in Chapter 4, they can focus on one area of their collection. They can also be selective in how they link outwards by making use of the LOD Cloud, meaning that a lot of the browsing ability does not need to be facilitated by the Archive itself. The collaborative LD approach could let people outside the Archive get involved to contribute their expertise to help further place the collection in a broader context. Since the Archive relies on Fanshawe's notes from the second half of last century, collaborative LD could help bring a modern understanding to Fanshawe's research. Collaborative LD could help the Archive present their resources accurately, precisely *and* colloquially.

5.3.Beyond Linked Data

For a small institution like the Fanshawe Archive there is always an element of playing catch-up to larger ones. The field of LD, where boundaries are always being pushed, is no exception. PCL (2019) discussed the difference in scale between national libraries and smaller libraries or archives, where national libraries are able, and expected, to lead the trend towards new technologies and practices in a way that smaller libraries can't. This is supported by Pennington and Cagnazzo (2019, p.3), who argue that national libraries are well placed to "lead the development of technological standards for record formats, data exchange, and interoperability protocols" due to their proximity to the government and subsequent ability to influence national policies. As small archives and libraries are only just coming to terms with LD and frantically trying to digitise their collections, national libraries and LD researchers are already exploring ways that LD can be combined with other technologies.

The anonymous national library's recent LD project combined LD with the Music Encoding Initiative (MEI), which allowed their external programmers to create further links within their musical material by encoding all the related sources to point to particular places (ANL, 2019). The technology allowed them to, say, link bar 4 in the composer's sketch to bar 4 in the conductor's annotated score, then encode both to a precise time stamp for bar 4 in performance video. The combination of MEI with LD is called MELD (Music Encoding and Linked Data) and could make a piece of music meaningful to computers the way RDF makes documents meaningful (Weigl and Page, 2015). While less relevant to the Fanshawe Archive, since their collection doesn't contain written music, the encoding ability could be a fantastic resource for libraries that carry written music. It would be unfeasible for the Archive, but MELD technology could theoretically encode Fanshawe's compositions that include quotations from his sound recordings (*African Sanctus* or *Pacific Odyssey*) to point to where those quotations come from, like the 'Bwala Dancers' or 'Call to Prayer'.

Another example of where LD technology could go, is the argument that LD can be combined with Social Tagging to make it more user-friendly (Pennington and Spiteri (eds.), 2019). Beyond that, Pennington (2019) hypothesises that, in the future, music clips and photographs could be used as non-textual social tags and suggests that if these tags had URIs they could be used as the predicate in the RDF triple. This could reduce language barriers for LD users, which is particularly pertinent in the field of World Music.

The extra steps described above demonstrate where LD could head in the future, particularly in the musical field, but it's unlikely that the Fanshawe Archive or other smaller archives could participate in these further steps any time soon. But, as technology and innovations progress, it might be a possibility in the future.

CHAPTER 6 – RECOMMENDATIONS AND CONCLUSIONS

6.1.Fanshawe Archive Recommendations

I agree with the Fanshawe Archive that they are not currently able to implement Linked Data, since their focus is on more pressing projects. However, they could make their catalogue LD Ready by choosing software that uses permalinks, visualising their Excel spreadsheet in terms of RDF structures and, where possible, trying to include hyperlinks to resources outside the Fanshawe Archive that can be more quickly LD coded in the future.

Once the Save Our Sounds project has digitised the last tapes in the Africa collection, the Archive could look towards creating an LD pilot with a small area of the collection, perhaps employing students or fixed-term contractors. If the archive makes it clear that it is potentially interested in participating in the LD projects of other cultural heritage institutions, this could help reduce the burden of cost, online storage, and time constraints.

The interconnected nature of the Fanshawe collection, with its journals and photographs that fit so well with the sound recordings, demonstrated the potential value of an internal RDF structure that could help computers understand the semantic relationships between items in the collection. Even if the Fanshawe Archive just focused on this, it would push them to 4 stars on the Berners-Lee LD development scheme. It could also make them a more appealing candidate for joining future LD projects and being linked to by external LD systems.

6.2.Wider Context Recommendations

For LD to work best, the whole web should ideally be using it. Not everyone will, and many old sites are no longer maintained at all. However, given their attention to detail, metadata understanding, pre-existing communities (like CILIP or IAML), large collections of resources, and general advocacy of open access to knowledge, libraries are well suited to lead the LD revolution even if other sites are slower to follow.

Librarians are, however, not necessarily computing experts, so in the long run, if LD is to be used widely in libraries and other cultural heritage institutions, I recommend the creation of a librarian-friendly LD OPAC system or at least a set of LD instructions specifically for librarians (i.e. written for an audience with a humanities background).

The isolation of small libraries means that professional bodies like CILIP or, in the case of music libraries, IAML, are vital resources for such librarians. If an institution like IAML were to facilitate the provision of resources to create a World Music Ontology for World Music collections to use, that

would make the creation of LD in that field much easier for small collections. Since LD is still in its infancy, if IAML, MLA or similar were to take on this kind of project now, they would be in a good position to standardise term usage and streamline the metadata of World Music from the beginning.

The Fanshawe Archive's collection would be a valuable resource to anyone trying to create a World Music Ontology, not just because of the vast number of different recordings from all over Africa and the Pacific, but also because of the extensive notes in the accompanying journals, rich with detail on the recordings which, as ANL (2019) pointed out, is not always the case – particularly with historical World Music collections.

Meanwhile, if an institution like CILIP, perhaps with the help of LD experts like OXLOD and/or the national libraries that are already implementing LD, were to make a concerted effort to fund, support and sponsor the creation of LD within smaller institutions, smaller institutions would be better equipped to keep up with LD. If those projects were properly documented and described, other smaller institutions would then be able to follow the project for their own libraries.

Although not a library, and therefore not within CILIP's jurisdiction, the Fanshawe Collection would be a good candidate for such an enterprise. The breadth of the collection, with complementary sound recordings, photos and journals, would give opportunities to link up resources inside and outside the collection. Also, by its nature, the collection is finite and not growing, meaning there is a definite endpoint. The collection is also split by geographical area, meaning that the project can be as wide or as small as would suit the researchers. Such a project would in turn help the Fanshawe Archive by spotlighting it for potentially interested parties to find more easily.

6.3.Conclusions

The project aimed to determine how LD could be used by the David Fanshawe World Music archive to make their collection more accessible online by answering the sub-questions laid out in the introduction.

Interviews with the Fanshawe Archive demonstrated that their current main goals are to create an online catalogue and digitise their collection to help preserve Fanshawe's legacy and transfer it to a more flexible format. They also hope to find somewhere to house the physical collection. Their current patronage is limited, but they are working to extend it by reaching out to higher education institutions and other potentially interested parties. They currently have no online presence but have plans to create a website that will alert people to the Archive's existence. Although they feel they don't have the necessary resources to devote to such a project, they would be in favour of LD if it made the collection more accessible online.

It has been shown through the literature and interview with the anonymous national library that LD can raise the online profile of a collection and is helpful and perhaps necessary for libraries and other cultural heritage institutions. The literature shows that LD has frequently been used in the field of music, and there are models such as Berners-Lee's (2009) 5-star development scheme and four best practice guidelines, and Pennington and Caganzzo's (2019) suggestions for best practice for Linked Library Data. But it is also widely acknowledged that smaller institutions are more limited in funding, time and sometimes technology.

The Fanshawe Archive has been carrying out its digitisation and cataloguing mostly alone, although they will collaborate with the Save Our Sounds project to complete the Africa tapes. Their catalogue is in its early stages, but they have created a template of how a user interface might look and plan to input the catalogue into an Excel spreadsheet. While they have access to resources that allow them to carry out these two projects, the necessary technical jump makes LD less realistically feasible, although everyone involved with the Archive is extremely dedicated and skilful in their fields. FPT was particularly enthusiastic about computers, but, at least currently, creating a full LD catalogue probably cannot be carried out by a single person. The anonymous national library outsourced most of the LD side of their project to an external team of experts.

The above recommendations detail how the Fanshawe Archive might still participate in LD and prepare for becoming more involved in the future, particularly through collaboration with libraries, wider bodies, technological engineers or other cultural heritage institutions' LD projects. But even if the Fanshawe Archive itself never decides to adopt an LD approach, the discussions in this dissertation about how the David Fanshawe World Music Archive could use LD to make its collection more accessible online show how LD might be used, thought about and worked towards in general. Hopefully this can guide and inform the wider library community on how they might approach LD themselves, particularly those interested in the contexts of traditional music and small archives.

REFERENCES

Andrews, P. (2003) 'Cecilia: towards a map of the music resource of the UK and Ireland', in Turbit, R. (ed.) *Music Librarianship in the United Kingdom: Fifty Years of the United Kingdom Branch of the International Association of Music Libraries, Archives and Documentations Centres*. Aldershot: Ashgate Publishing Limited, pp. 200-208.

ANL (Anonymous national library) (2019) Interviewed by Kirsty Morgan. 8 July.

Bartolome, S. J. Mukuna, K. and Oehrle, E. (2010) 'Music of Sub-Saharan Africa' in Anderson, W. M. and Campbell, P. S. (eds.) *Multicultural Perspectives in Music Education Volume 1*, Plymouth: Rowman & Littlefield Education, pp. 7-44.

Bates, M. E. (2013) 'Of Hashtags and Descriptors'. *Online Searcher*, 37(3), p. 80.

Batthini, G. (2014) 'Information Technology Skills for Library Professionals', *SSRN Electronic Journal*. Available at: https://www.researchgate.net/publication/272248175 (Accessed: 16th August 2019).

Berners-Lee, T. (2009) *Linked Data*. Available at: https://www.w3.org/DesignIssues/LinkedData.html (Accessed: 17th January 2019).

Bizer, C. Heath, T. and Berners-Lee, T. (2009) 'Linked Data--the story so far'. *International Journal on Semantic Web and Information Systems*, 5(3), pp. 1-22.

Brandt, J.C. (2016). 'Establishing Aboriginal Presence in the Museum Sector', in Callison, C. Roy, L. and LeCheminant, G. A. (eds.) *Indigenous Notions of Ownership and Libraries, Archives and Museums*. Berlin: De Gruyter Saur, pp. 210-228.

The British Library Board (2009) *Legal and Ethical Usage*. Available at: https://sounds.bl.uk/Information/Legal-And-Ethical-Usage (Accessed: 8th March 2019).

Cagnazzo, L. (2019) 'Tagging the semantic web: combining Web 2.0 and Web 3.0', in Pennington, D. and Spiteri, L. (eds.) *Social Tagging in a Linked Data Environment*. London: Facet Publishing. pp. 11-37.

Chang, H. and Iyer, H. (2012) 'Trends in twitter hashtag applications: Design features for value-added dimensions to future library catalogues'. *Library Trends,* 61(1), pp. 248-258.

Chen, S. S. (2016). 'A Holistic Perspective on Indigenous Digital Libraries in Taiwan', in Callison, C. Roy, L. and LeCheminant, G. A. (eds.) *Indigenous Notions of Ownership and Libraries, Archives and Museums*. Berlin: De Gruyter Saur, pp. 311-324.

Chisita, C. T. Rusero, A. M. and Shoko, M. (2016). 'Leveraging Memory Institutions to Preserve Indigenous Knowledge in the Knowledge Age', in Callison, C. Roy, L. and LeCheminant, G. A. (eds.) *Indigenous Notions of Ownership and Libraries, Archives and Museums*. Berlin: De Gruyter Saur, pp. 273-285.

Connelly, K. *et al.* (2019) *Digital Transformation for UK Public Libraries: Five Approaches to a 'Single Digital Presence'*. Available at: https://bit.ly/2I1p9aR (Accessed: 12th August 2019).

Coyle, K. (2010) 'Chapter 1: Library data in a modern context'. *Library Technology Reports*, 46(1), p. 5(9).

Dempsey, L. (2000) 'Scientific, Industrial, and Cultural Heritage: A Shared Approach'. *Ariadne*, 22, Available at: http://www.ariadne.ac.uk/issue22/dempsey/ (Accessed: 29th May 2019).

Dobson, M. (2019) '#FandomCommunication: how online fandom utilises tagging and folksonomy', in Pennington, D. and Spiteri, L. (eds.) *Social Tagging in a Linked Data Environment*. London: Facet Publishing. pp. 131-150.

Ellingham M. Duane, O. and McConnachie, J. (eds.) (2000) *World Music Volume 2: Latin & North America, Caribbean, India, Asia and Pacific,* London: Rough Guides Ltd.

Fanshawe, D. (1973) *Treasures from Uganda* [Radio Programme] 26th October. Recording held at David Fanshawe World Music Archive, Malvern (Accessed: 27th June 2019).

Fanshawe, D. (1987) *Musical Mariner: Pacific Journey* [DVD]. Sydney: Lucky Country Productions.

Fanshawe, D. (1998) *Spirit of Melanesia* [CD]. Badminton: Saydisc Records.

Fanshawe, D. (2002) *Music of the South Pacific: Recordings by David Fanshawe* [CD]. West Sussex: ARC Music Productions Int. Ltd.

Fanshawe, D. *et al.* (2006) *African Sanctus revisited: David Fanshawe's African journey* [DVD]. Marlborough: Fanshawe One World Music.

Fay, R. M. and Sauers, M. P. (2012) *Semantic Web Technologies and Social Searching for Librarians*, London: Facet Publishing.

The Folk Shop (2019) *Instrument Descriptions: F-M* Available at: https://www.thefolkshop.com/instrument-descriptions-f-m (Accessed: 14th July 2019).

FM (Fanshawe Manager) (2019) Interviewed by Kirsty Morgan. 27 June.

FPT (Fanshawe Part-time) (2019) Interviewed by Kirsty Morgan. 25 June.

Francis, H. *et al.* (2016). 'Accessing Sound at Libraries, Archives, and Museums', in Callison, C. Roy, L. and LeCheminant, G. A. (eds.) *Indigenous Notions of Ownership and Libraries, Archives and Museums*. Berlin: De Gruyter Saur, pp. 344-368.

FT(D) (Fanshawe Trustee (Digital)) (2019) Interviewed by Kirsty Morgan. 26 June.

FT(L) (Fanshawe Trustee (Librarian)) (2019) Interviewed by Kirsty Morgan. 26 June.

Futornick, M. (2019) *Performed Music Ontology*. Available at: https://wiki.duraspace.org/display/LD4P/Performed+Music+Ontology (Accessed: 12th August 2019).

GCH (Gerald Coke Handel Library) (2019) Interviewed by Kirsty Morgan. 21 June.

Gracy, K. F. Xeng, M. L. and Skirvin, L. (2013) 'Exploring Methods to Improve Access to Music Resources by Aligning Library Data with Linked Data: A Report of Methodologies and Preliminary Findings'. *Journal of the American Society for Information Science and Technology*, 64(10), pp. 2078-2099.

Harpring, P. (2010) *Introduction to Controlled Vocabularies*, Los Angeles: Getty Publications.

Hausenblas, M. (2012) *5* Open Data*. Available at: https://5stardata.info/en/ (Accessed: 30th July 2019).

Hill, R. (2017) *Free Data Services*. Available at: http://www.bl.uk/bibliographic/datafree.html (Accessed: 10th March 2019).

Hogg, K. (2015) 'David and Goliath: How to be a Small Library in a Big Digital World'. *Brio*, 52(2), pp. 46-56.

Hudson, R. (2015) *Visualizing Cultural Heritage: Linked Open Data and the Carnegie Hall Archives p.1*. Available at: https://www.pilsudski.org/pl/nowosci/blog/529 (Accessed: 22nd July 2019).

Kelly, E. J. (2015) 'Linked Data and Music: Current Projects and Opportunities'. *The Indexer*, 33(1), pp. 2-7.

Kubik, G. (2010) *Theory of African Music: Volume 1*, Chicago: The University of Chicago Press.

Leimbach, C. (1987) 'David Fanshawe: Musical Mariner', *Geo: Australia's Geographical Magazine*, 9(2), pp. 24-37.

McCallum, S. H. (2016) 'BIBFRAME and Linked Data for Libraries', in Jones. E. and Seikel, M. (eds.) *Linked Data for Cultural Heritage*. London: Facet Publishing, pp. 105-124.

McCann, H. S., Pulsifer, P. L. and Behe, C. (2016). 'Sharing and Preserving Indigenous Knowledge of the Arctic Using Information and Communications Technology', in Callison, C. Roy, L. and LeCheminant, G. A. (eds.) *Indigenous Notions of Ownership and Libraries, Archives and Museums*. Berlin: De Gruyter Saur, pp. 126-144.

McCarthy, M. Date, C. and Fanshawe, J. (2018). *A Special Collection: The David Fanshawe World Music Archive*. Malvern: David Fanshawe World Music Archive Trust.

McCrae, J. P. et al. (2019) *The Linked Open Data Cloud*. Available at: https://lod-cloud.net/ (Accessed: 7th March 2019).

OCLC Developer Network Team (2019) *Fanshawe, David, 1942 2010 Virtual International Authority File*. Available at: http://viaf.org/viaf/85318078 (Accessed: 31st July 2019).

O'Dell, A. J. (2016) 'Authority Control for the Web: Integrating Library Practice with Linked Data', in Jones. E. and Seikel, M. (eds.) *Linked Data for Cultural Heritage*. London: Facet Publishing, pp. 41-54.

Page, K. (2017) 'Representation of *Ring* Annotation Data', in Cafiero, F. (2017) 5th July. Available at: https://twitter.com/F_Cafiero/status/882552423239122944/photo/1 (Accessed: 8th July 2019).

Page, K. *et al.* (2017) 'Linked Data Publication of Live Music Archives and Analyses' in Hutchison, D. *et al.* (eds.) *The Semantic Web – ISWC 2017*. Cham: Springer Nature, pp. 29-37.

Papakhian, A. R. (2000) 'Cataloguing', in Griscom, R. (ed.) *Music Librarianship at the Turn of the Century*. Lanham: Scarecrow Press Ltd. pp. 19-28.

Pasaribu, I. (2016). 'Indigenous Digital Oral History', in Callison, C. Roy, L. and LeCheminant, G. A. (eds.) *Indigenous Notions of Ownership and Libraries, Archives and Museums*. Berlin: De Gruyter Saur, pp. 325-343.

PCL (Piping Centre Library) (2019) Interviewed by Kirsty Morgan. 19 July.

Pennington, D. (2016) 'Demystifying Linked Data: are you ready for what's next?', *CILIP Update, 2016(July/August)*, pp. 34-46.

Pennington, D. (2019) 'Keys to their own voices: social tags for a dementia ontology as a human right', in Pennington, D. and Spiteri, L. (eds.) *Social Tagging in a Linked Data Environment*. London: Facet Publishing. pp. 151-168.

Pennington, D. and Cagnazzo, L. (2019) 'Connecting the silos: Implementations and perceptions of linked data across European libraries'. *Journal of Documentation*, 75(3), pp. 643-666.

Pennington, D. and Spiteri, L. (eds.) (2019) *Social Tagging in a Linked Data Environment*. London: Facet Publishing.

Pugin, L. (2015) 'The Challenge of Data in Digital Musicology'. *Frontiers in Digital Humanities*. Available at: https://www.frontiersin.org/articles/10.3389/fdigh.2015.00004/full (Accessed: 16th August, 2019).

Qiolevu, L. (2015) 'A New Journey Begins', *Fiji Sun*. Available at: https://fijisun.com.fj/2015/08/25/a-new-journey-begins/ (Accessed: 29th July 2019).

Raimond, A. Y. *et al*. (2013) *Specification*. Available at: http://musicontology.com/specification/ (Accessed: 9th March 2019).

Redfern, B. (1979) *Organising Music in Libraries Volume 2: Cataloguing*, London: Clive Bingley Ltd.

Richards, L. (2006) *Handling Qualitative Data: A Practical Guide*, London: Sage Publications Ltd.

Riley-Huff, D. A. and Rholes, J. M. (2011) 'Librarians and Technology Skill Acquisition: Issues and Perspectives'. *Information Technology and Libraries*, 30(3), pp. 129-140.

Roberts, L. and Cohen, S. (2013) 'Unauthorising popular music heritage: outline of a critical framework'. *International Journal of Heritage Studies*, 20(3), pp. 241-261.

Rose-Steel, T. and Turnator, E. (2016) 'Medieval Music in Linked Open Data: A Case Study on Linking Medieval Motets'. *International Journal of Humanities & Arts Computing: A Journal of Digital Humanities*, 10(1), pp.36-50.

Roy, L. (2016). 'Who is Indigenous?', in Callison, C. Roy, L. and LeCheminant, G. A. (eds.) *Indigenous Notions of Ownership and Libraries, Archives and Museums*. Berlin: De Gruyter Saur, pp. 7-24.

Ruddock, B. (2019) *National Bibliographic Knowledgebase*. Available at: https://www.jisc.ac.uk/rd/projects/national-bibliographic-knowledgebase (Accessed: 24th July 2019).

Ruddock, B. and Stevenson, J. (2011) 'Creating Linked Open Data for Library and Archive Descriptions'. *Multimedia Information & Technology*, 37(4), pp.19-20.

Save Our Sounds (2019). Available at: https://www.bl.uk/projects/save-our-sounds (Accessed: 2nd June 2019).

Shotton, D. (2013) *Libraries and linked data #1: What are linked data?* Available at: https://semanticpublishing.wordpress.com/2013/03/01/lld1-what-are-linked-data/ (Accessed: 24th May 2019).

Spiteri, L. F. (2019) 'Hashtags and library discovery systems', in Pennington, D. and Spiteri, L. (eds.) *Social Tagging in a Linked Data Environment*. London: Facet Publishing. pp. 89-108.

Sporny, M. (2012) *What is Linked Data?* Available at: https://m.youtube.com/watch?v=4x_xzT5eF5Q (Accessed: 28th July 2019).

Sweeney, P. (2011) *The World in his Ear* [Radio Programme] 4th October. Available at: https://www.bbc.co.uk/programmes/b015ck9b (Accessed: 29th June 2019).

Syn, S. Y. (2019) 'Social tags for linked data with Resource Description Framework (RDF)', in Pennington, D. and Spiteri, L. (eds.) *Social Tagging in a Linked Data Environment*. London: Facet Publishing. pp. 39-58.

Thomas, G. (2011) *How to do your Research Project*, London: Sage Publications Ltd.

Thorsen, H. K. and Pattuelli, M. C. (2016) 'Linked Open Data and the Cultural Heritage Landscape', in ed. Jones. E. and Seikel, M. *Linked Data for Cultural Heritage*. London: Facet Publishing, pp. 1-22.

Transforming Musicology (2019). Available at: https://tm.web.ox.ac.uk/home (Accessed: 8th July 2019).

Van Hooland, S. and Verborgh, R. (2014) *Linked Data for Libraries, Archives and Museums: How to Clean, Link and Publish Your Metadata*, Chicago: ALA Neal-Schuman.

Villanueva, C.B. (2016). 'The University of the Philippines Baguio Cordillera Studies Collection Library', in Callison, C. Roy, L. and LeCheminant, G. A. (eds.) *Indigenous Notions of Ownership and Libraries, Archives and Museums*. Berlin: De Gruyter Saur, pp. 286-310.

VWM (Vaughan Williams Memorial Library) (2019) Interviewed by Kirsty Morgan. 3 July.

Weigl, D. and Page, K. (2015) *Music Encoding and Linked Data (MELD)* Available at: http://www.semanticaudio.ac.uk/demonstrators/24-music-encoding-and-linked-data-meld/ (Accessed: 2nd August 2019).

Weissenberger, L. (2017) 'Stories, Songs, Steps, and Tunes: A Linked Data Ontology for Irish Traditional Music and Dance' *International Society for Knowledge Organization, UK/Ireland Chapter, Knowledge Organisation: What's the Story?* London, United Kingdom, 11-12 September 2017. Available at: https://www.researchgate.net/publication/320197501 (Accessed: 2nd June 2019).

Weissenberger, L. (2018) ‘The Litmus Linked Data Project at the Irish Traditional Music Archive’. *Brio*, 55(1), pp. 52-57.

W3C (2014) *Best Practices for Publishing Linked Data*. Available at: https://dvcs.w3.org/hg/gld/raw-file/default/bp/index.html (Accessed: 28th July 2019).

What is WorldCat? (2019). Available at: https://www.worldcat.org/whatis/default.jsp (Accessed: 28th July 2019).

Zhao, L. and Ichise, R. (2014) ‘Ontology Integration for Linked Data’. *Journal on Data Semantics*, 3(4), pp. 237-254.

APPENDIX A – BRIEFING NOTE, PRIVACY POLICY, CONSENT FORM AND INTERVIEW QUESTIONS

An example of one of the documents sent in advance of the interview. It contains a research brief, privacy statement, consent form and list of interview topics. This one was sent to the Fanshawe Manager (FM).

Briefing Note, Privacy Policy and Consent Form

Brief Overview

Thank you for agreeing to be part of my dissertation, "Linked Data in Small Archives: A Case Study of the David Fanshawe World Music Archive". As technology progresses, it is important to find ways to keep relatively small but significant archives relevant and accessible; otherwise, they're in danger of disappearing and their collections will be lost. Consequently, this project aims to carry out a case study of the Fanshawe Archive to determine whether Linked Data could be implemented to make the archive's collection more accessible online. This includes collecting information about resources, processes and technical requirements. I will also collect information about how cultural heritage institutions collaborate and help each other, because this is relevant to determining the feasibility of big Linked Data projects. The dissertation will also look at the specific challenges of using Linked Data in the music discipline, since music cataloguing has extra requirements that traditional literature doesn't. As part of the research, I am conducting interviews with staff from the Fanshawe Archive and from four other libraries to place the Fanshawe Archive into a wider context and to gather information about previous projects to make libraries more accessible online (particularly, but not exclusively, within the Linked Data sphere).

Below you will find a list of the questions that I plan to cover. This will be a semi-structured interview, so there may be follow-up questions based on the answers you give, and I may ask you to elaborate on specific areas of interest, but the order and content will be what you see here.

GDPR Ethics

Your interviews will be used for the purposes of my dissertation. Transcripts of the interviews will be created to aid analysis, but not included in full in the dissertation. The audio of the interview will be recorded to make the transcription process easier and more accurate, unless requested otherwise. Whether the interview is recorded or not, you will be sent a copy of the final interview transcript/notes by email.

While I will not be collecting personal or sensitive data, please be aware that, especially for smaller archives and libraries, it may be possible to identify you personally from your job role and/or institution. For this reason, I have provided the option to anonymise either of these areas, if requested.

After I have received my dissertation mark and graduated in November 2019, I will delete the audio files and the interview transcripts from any devices I have them on (using recycle bin and emptying it). However, be aware that the transcript will have been sent to you for approval via email, so I will delete my emails, but for the data to be removed completely you will need to delete your emails, too. They will also be sent between my dissertation supervisor and myself, and I will request that she delete her emails too.

Quotes from the interview will be findable in electronic versions of my dissertation. It will be held by the university, might be put on StrathPrints (a university research portal) and may be sent to other relevant institutions, for example IAML. I will personally keep one electronic copy of my dissertation in case I need it in future.
Under GDPR, you have the right to request access to all of the information held on you and request to know how it is being processed and/or that it be deleted. Until the 16th of August (my dissertation submission date), the person controlling that data is me, you can contact me at kirsty.morgan.2018@uni.strath.ac.uk. After that, please contact the University of Strathclyde via the ethic committee at the address ethics@cis.strath.ac.uk. If they are unable to comply with your request, they will direct you to the appropriate data controller.

Consent Form

Please complete the consent form and return to me at kirsty.morgan.2018@uni.strath.ac.uk

Your name: _____
Date: _____

Please indicate Y or N whether you consent to:

Being interviewed with the questions below	
The audio of the interview being recorded to make it easier to transcribe later	
The name of your institution being identified	
Your job role being identified	
Quotes from the interview being used in the body of the dissertation	

Please note that you are free to withdraw from the interview at any time without providing a reason. You are also free to decline to answer any of the questions without giving a reason.

Contact Details
The Institution: University of Strathclyde
The Researcher: Kirsty Morgan, Information and Library Studies Masters student, kirsty.morgan.2018@uni.strath.ac.uk
Dissertation Supervisor: Dr. Diane M. Rasmussen Pennington, Senior Lecturer in Information Science, diane.pennington@strath.ac.uk, +44 (0)141 548 3900
Departmental Ethics Committee: Prof. Ian Ruthven, Dr. Lisa McCann, and Dr Marc Roper, ethics@cis.strath.ac.uk

Interview with Fanshawe Archive Manager

A. INTRODUCTION
A.1.1. What is your role within the Fanshawe Archive?
A.2. Linked Data Understanding
A.2.1. How much, if anything, do you know about Linked Data?

B. ABOUT THE INSTITUTION
B.1. Vision for Institution
B.1.1. How would you describe the goals, intentions and purpose of the Fanshawe Archive?
B.2. Patron Questions
B.2.1. What kind of patrons does the Fanshawe Archive cater to?
B.2.2. What do you do to accommodate your current patrons?
B.2.3. Is there anything you do to attract new patrons?
B.3. Making Institution Accessible
B.3.1. Does the Fanshawe Archive have an online presence?
B.3.2. Can you talk about that?
B.3.3. Has anything been done or put in place to increase online traffic to your site?

C. PROJECT REQUIREMENTS
C.1. Resource Questions
C.1.1. What resources does the Fanshawe Archive currently have access to?
C.1.2. Do you think the Fanshawe Archive has or would be able to gain access to the resources required to undertake a Linked Data project?
C.2. Current Tech Situation
C.2.1. Can we talk about the current technological setup in the Fanshawe Archive? What sorts of technology does the Fanshawe Archive currently have?
C.2.2. You are currently working on a project to digitise the Fanshawe Archive's collection; can you talk a bit about that process?
C.2.3. What is the technology expertise of the staff?
C.2.4. Can you tell me about the Fanshawe Archive's current catalogue?
C.2.5. Are there any challenges you have experienced when cataloguing/describing indigenous/music sources?

D. COLLABORATION BETWEEN CULTURAL HERITAGE INSTITUTIONS
D.1. Getting Support
D.1.1. Do you currently receive any support from outside sources for big projects? What kind?
D.1.2. Would you want (extra) support? What kind would you want?
D.1.3. How would you go about getting it?
D.2. Giving Support
D.2.1. Does the Fanshawe Archive collaborate with or offer assistance to other (smaller) institutions for undertaking larger projects?
D.2.2. In what way?
D.2.3. Would it be something you could see the Fanshawe Archive doing in the future?

E. THE FUTURE AND OTHER CONSIDERATIONS
E.1. Ethics of Using Indigenous Sources
E.1.1. Some ethical considerations are raised when dealing with world music and indigenous music. Do you have any thoughts on that side of things?
E.2. Future Hopes for Linked Data
E.2.1. If you were to undertake a Linked Data project, what would you hope to get out of it?
E.2.2. Are there any potential drawbacks that you can see from implementing a Linked Data project?

E.2.3. As you understand it, what would Linked Data allow the Fanshawe Archive to do or achieve that isn't currently possible?

E.3. Any Other Comments

E.3.1. Is there anything else you feel you'd like to talk about?

APPENDIX B – DIAGRAMS

Fig.5.3. Example Concept Map for the Madinda Recording (FAC 44) © Images copyright David Fanshawe

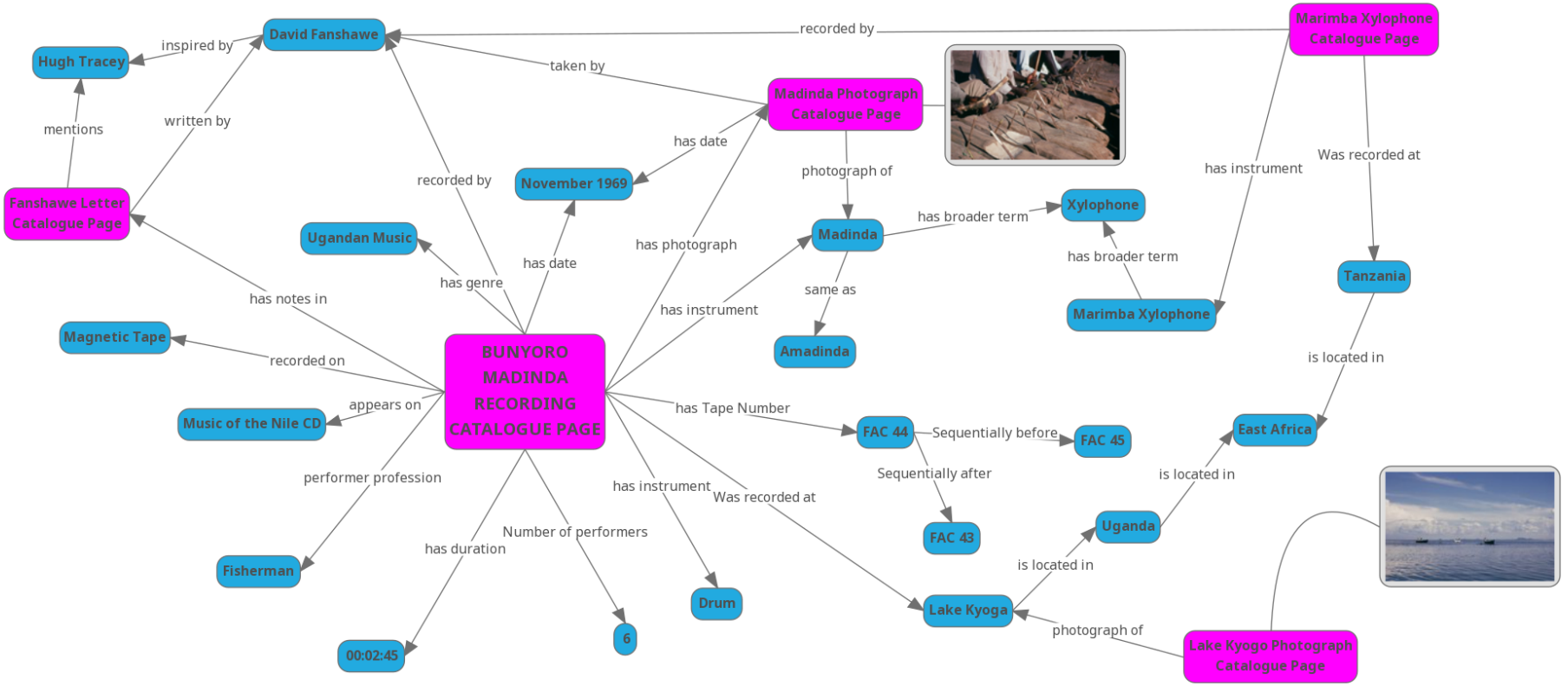

Fig.5.4. Example Graph for the Madinda Recording (FAC 44)

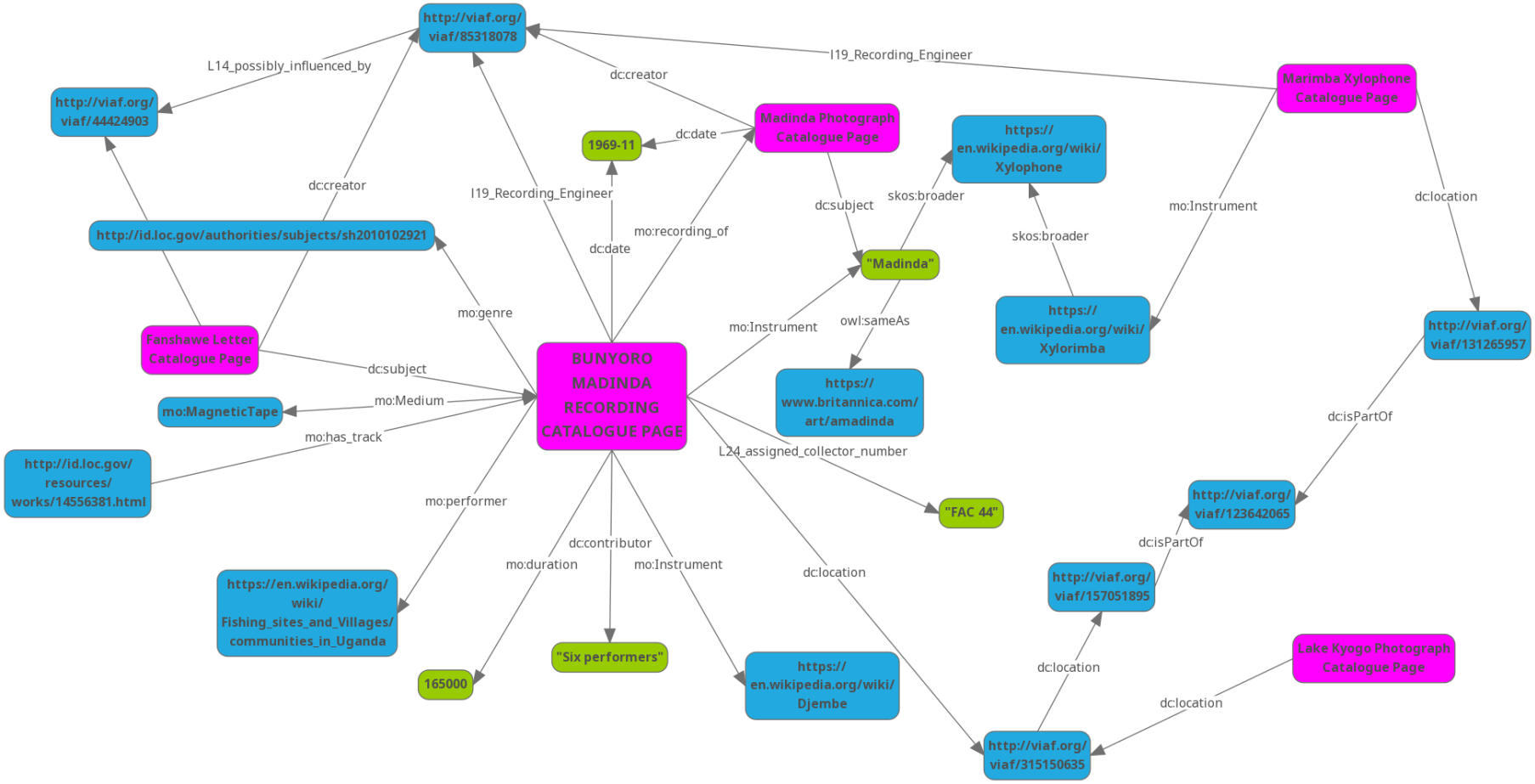

Fig.5.5. Example cut-down Graph for the Madinda Recording

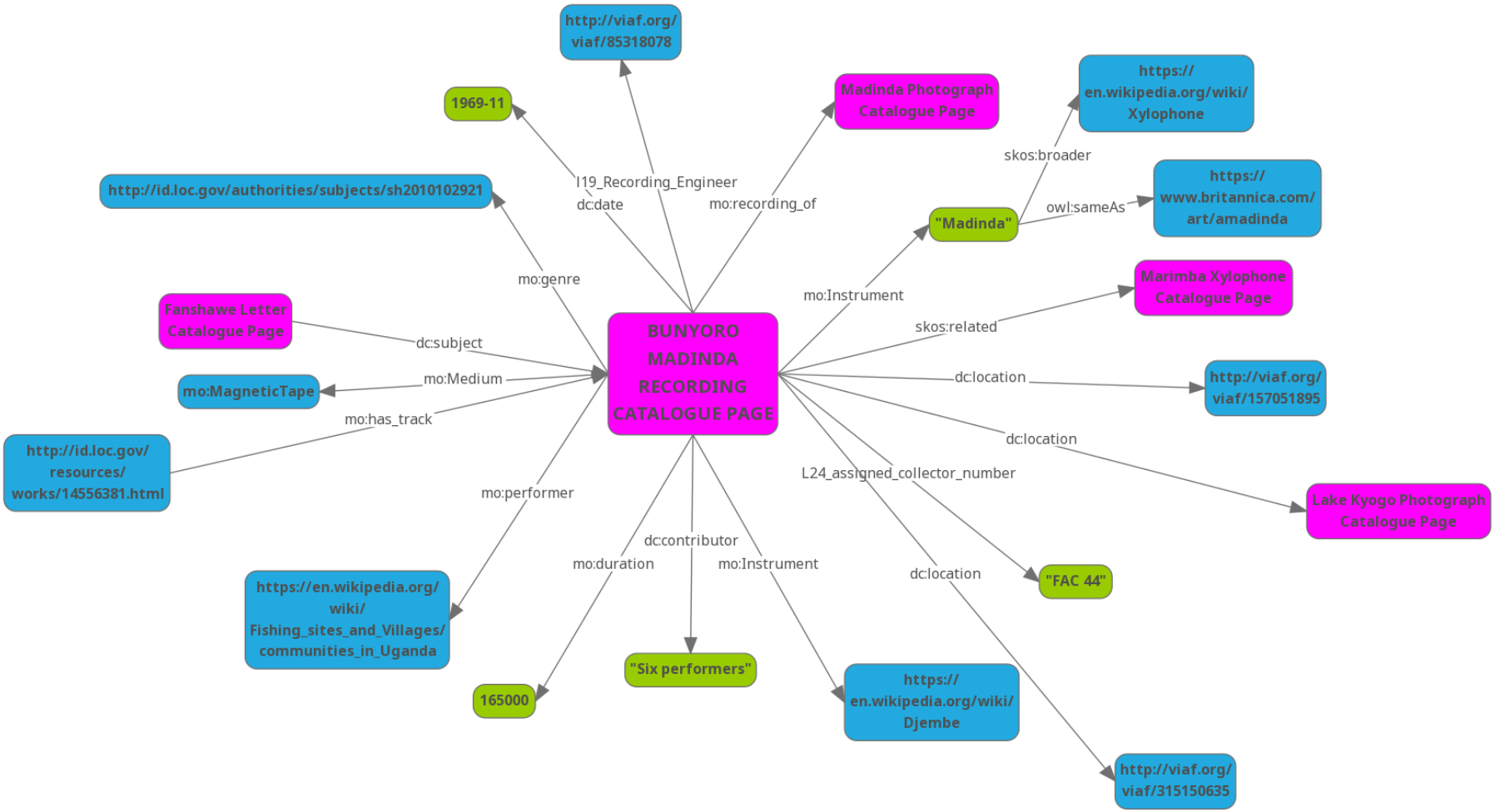

Fig.5.8. Example external links for the Wagi Brothers Recording (FPC 1389)