Software Tools for Indigenous Knowledge Management

Bevan Raymond Koopman

October 25, 2002

University of Queensland

Software Tools for Indigenous Knowledge Management

Bachelor of Information Technology Honours Thesis,
School of Information Technology and Electrical Engineering

Supervisors:
Dr Jane Hunter, CRC for Distributed Systems Technology
Prof. Simon Kaplan, University of Queensland

by

Bevan Raymond Koopman

Keywords: Indigenous, Knowledge, Rights, Management, Software, Multimedia

Acknowledgments

I would like to specially thank Dr. Jane Hunter for giving me the opportunity to work on this project. For the wealth of experience and knowledge that I gained, not just in the field of multimedia and metadata, but in computer science research in general, I am very grateful.

The Distributed Systems Technology Centre provided me with excellent working conditions and an exciting and hospitable environment. I am proud to be affiliated with them.

Thank you to Angelina Russo and the three Information Environments students who provided a very professional graphical interface.

Finally, I would like to thank my family and Jennifer Dale for their continued support throughout the last year.

This thesis acknowledges the plight of the Aboriginal people, who's culture and communities have suffered as a result of colonisation, western law and globalisation.

Bevan Koopman, Brisbane 2002

Abstract

Indigenous communities are beginning to realise the potential benefits which digital technologies can offer with regard to the documentation and preservation of their histories and cultures. However they are also coming to understand the opportunities for misuse and misappropriation of their knowledge which may accompany digitisation. In this thesis we describe a set of open source software tools which have been designed to enable Indigenous communities to protect their unique cultures and knowledge which are being preserved through digitisation. The software tools described here enable authorised members of communities to: define and control the rights, accessibility and reuse of their digital resources; uphold tribal customary laws pertaining to secret/sacred knowledge or objects; prevent the misuse of Indigenous heritage in culturally inappropriate or insensitive ways; ensure proper attribution; and finally to enable communities to describe their resources in their own words. Hopefully the deployment of such tools will contribute to the self-determination and empowerment of Indigenous communities through the revitalisation of their cultures and knowledge which have been eroded by colonisation, western laws, western cultures and globalisation.

Contents

List of Figures

Chapter 1

Introduction

In recent years Indigenous communities have recognised the importance of documenting and sharing their heritage and history. This comes at a time of exponential growth of the Internet and the widespread use of multimedia technologies. This has triggered demand for new copyright protection mechanisms to manage usage and ensure proper compensation. A number of new technologies have been developed to govern the use of online resources. These include XML-based markup languages such as XrML [1] and ODRL [2] for the expression of rights, as well as the standardisation of a rights expression language called for by the Moving Picture Expert Group (MPEG), within the MPEG-21 [3] Framework.

However these new technologies are based on modern intellectual property laws and often fail to meet the unique requirements of Indigenous knowledge. The concept of ownership in Indigenous cultures differs from that of the western notion, with Indigenous cultures having the notion of collective rather than individual ownership. Permission to access particular resources in Indigenous communities is often far more complex than in western communities. Factors such as a person's gender, status within a tribal unit, role and relationships to people being represented in the resource are all determining factors. These factors can differ vastly between communities, making it difficult to define a standard set of laws and protocols. As a result, current methods of expressing rights restrictions and enforcing access controls do not apply well to Indigenous knowledge.

The aim of this thesis is to evaluate existing international standards for content description, resource discovery and rights management. Then based on these standards and extensions implement a set of open source software tools for the management of Indigenous knowledge.

1.1 Objectives

This thesis investigates how modern information technologies can be used in the preservation of Indigenous knowledge. In cases where these technologies fail it will present extensions to allow them to be applied to the unique domain of Indigenous knowledge.

The specific goals of this project are to develop information technology tools which enable:

- Digital representation of Indigenous knowledge;
- Traditional owners to define ownership, dissemination, access and copyright permissions;
- Support for customary laws and protocols;
- Enforcement of access and usage constraints;
- Protection against misappropriation;
- Proper attribution and compensation to traditional owners;
- Preservation of Indigenous knowledge, cultures, history.

This project is being conducted with the Distributed Systems Technology Centre (DSTC) within the Multimedia Access for Enterprises Across Networks and Domains (MAENAD) research group. DSTC is talking to a number of Indigenous communities, discussing possible collaborative projects. The political nature of dealing with these different groups, and the sensitive nature of Indigenous knowledge has made this process slow and often difficult.

A more specific outline of the project's objectives are presented below.

1.1.1 Representation of Indigenous Knowledge in a Digital Form

Most Indigenous resources [4] held in museums, archives, cultural centres and Keeping Places are not currently in digital form. These resources (such as manuscripts, photographs and maps) are often stored in long-term archives, with some having deteriorated as a result of age and environmental factors. The knowledge they provide is only available to a select few who have permission to access them. Digitisation provides a way of both preserving these resources and making them more freely available.

Representing these resources in a digital form involves more than simply digitising them. It is important to also include information that describes the resource and defines access and usage

contraints, its metadata. An effective and standardised model of representing and distributing this metadata needs to be developed to give meaning to each digital item.

This model needs to encompass both descriptive and rights metadata and should be standardised and support interoperability. Any sensitive data needs to be stored in a way that ensures confidentiality and integrity.

1.1.2 User Definition of Rights

Due to the complexity and nature of Indigenous intellectual property laws it is important to allow traditional owners to define ownership, dissemination, access and copyright permissions. They should clearly and easily be able to specify how and by whom a particular resource may be used. In addition, there should be support for defining customary laws and protocols. These definitions should be represented in a standard portable form.

1.1.3 Enforcement of Access and Usage Constraints

It is not simply enough to define access and have usage constraints represented in a consistent manner. These restrictions need to be translated into a machine understandable representation and their semantics enforced. Software tools are needed to enforce access control on the basis of a user's profile and the rights expressions associated with the resource they wish to use.

1.1.4 Annotations

A vast amount of Indigenous knowledge does not exist in physical or digital form, but purely as knowledge in the minds of academics, tribal elders and people of remote communities. This knowledge is lost if the person dies or does not pass it on to others. As a result there is considerable need for people to be able to describe and annotate resources in their own words (or own language).

1.1.5 Search, Browse and Retrieval Facilities

As stated previously digitisation provides a way of making the knowledge represented by valuable Indigenous resources available to a wider audience. In addition to facilities which enable the description of a collection of resources, and govern how they should be used, a mechanism is also required for searching across the collection and viewing or browsing resources of interest.

As these collections typically involve a number of different multimedia types, users need to be presented with their search results in a meaningful way. This is done through the generation of multimedia presentations.

1.2 Scope

Chapter 2 presents background and related work. It reviews current knowledge management technologies, investigates the unique requirements of Indigenous knowledge and the limitations of modern copyright laws and IT mechanism when applied to Indigenous knowledge.

Chapter 3 contains a software requirements specification for the proposed set of software tools to be developed.

Chapter 4 describes the high level system architecture, the metadata model used and workflow of the software.

Chapter 5 describes the Metadata Editor / Generator application which forms the back-end, administrative and metadata input component of the system.

Chapter 6 describes the Search, Browse & Retrieval Interface which forms the front-end, search and presentation component of the system.

Chapter 7 describes the problems relating to this project, presents possible future work, and contains conclusions.

Chapter 2

Background & Related Work

This chapter provides readers with the necessary background to the material later presented in this thesis. The literature reviewed in this thesis was of two main types: work in the information technology area pertaining to knowledge management technologies with recent research in digital rights management; and non-technical literature in Indigenous knowledge concerned primarily with anthropological or copyright issues. Literature that incorporates these two areas is rare. This thesis aims to bring these two separate domains together.

In order to understand the area of digital knowledge management current technologies are reviewed. The unique requirements of the Indigenous knowledge domain are then presented so that the reader may understand how current knowledge management technologies fail when applied to Indigenous knowledge.

One of few pieces of literature that considers Indigenous knowledge in the information technology domain is then described. This literature provides the theoretical background for this thesis and is written by the supervisor of this thesis.

2.1 Current Knowledge Management Technologies

In this section current knowledge management technologies are considered. Some insight is given into how these technologies may be used for the management of Indigenous knowledge, while a later section considers their shortcomings in the Indigenous knowledge domain.

2.1.1 Describing Multimedia Content

Dublin Core

The Dublin Core Metadata Initiative [5] is an organisation promoting the widespread adoption of an interoperable metadata standard for simple resource discovery. Currently Dublin Core supports the following fields: *Title*, *Creator*, *Subject*, *Description*, *Publisher*, *Contributor*, *Date Created*, *Date Issued*, *Type*, *Identifier*, *Source*, *Language*, *Relation*, *Coverage*, *Rights*.

Dublin Core provides a standard way of representing the descriptive metadata associated with an Indigenous resource. Having a standardised way to describe content means that it will be discoverable by many search interfaces and is interoperable across many domains, thus allowing sharing and easy distribution of information and improved accessibility.

CIDOC CRM

The CIDOC object-oriented Conceptual Reference Model (CRM [6]) was developed by the ICOM/CIDOC Documentation Standards Group as an ontology for cultural heritage information. The CIDOC CRM is presented as an object-oriented extensible data model normally defined using a RDF Schema. It serves as a basis for mediation of cultural heritage information, allowing information exchange and integration of heterogenous resources between cultural institutions.

Hunter produced a paper [7] outlining how the CIDOC CRM ontology may be merged with the MPEG-7 ontology (for multimedia content description) to produce a single ontology for describing museum multimedia content. Hunter's proposed interoperable model, supporting both CIDOC CRM and MPEG-7 metadata could be used as the basis to provide a metadata model for describing Indigenous multimedia resources.

2.1.2 Rights Expression Languages

The vast increase in multimedia content providers and consumers, and the opportunity for exploitation and misuse, has led to the development of new technologies to govern how multimedia content is delivered and used.

The first step in ensuring the proper use of digital content is devising a standardised way of describing the usage of digital content. This has lead to a number of proposals from both rights and multimedia groups.

XrML

XrML (eXtensible rights Markup Language) [1] is a rights expression language designed by ContentGuard. It is used for specifying rights, fees and conditions for using digital content. It provides a set of structural and semantic tags defined by an XML Schema. These can be used to specify the rights metadata of digital content.

ODRL

The Open Digital Rights Language (ODRL [2]) is another rights expression language very similar to XrML. Its grammar also uses XML Schemas to specifcy syntax and semantics.

MPEG-21

The Moving Picture Expert Group is currently working on MPEG-21 [3], a Multimedia Delivery Framework standard. It aims to provide an infrastructure for the delivery and consumption of all content types, by different users, in different application domains. MPEG-21 has identified the need for a standardised machine-readable language to describe rights. It has defined the requirements for a Rights Data Dictionary and Rights Expression Language, and is currently developing these components.

2.1.3 Extensible Markup Language (XML)

XML Schemas

A relatively new addition to XML, XML Schemas [8][9][10] are similar to XML DTDs: they both describe the syntax and semantics of a XML document. XML Schemas differ in that they are written in XML themselves. This means that a schema can be navigated and manipulated programmatically just like a XML document. XML Schemas also provide support for namespaces, extensibility, data types, cardinality and inheritance. XML Schemas would be the tool of choice in describing a rights expression language that can be used for Indigenous rights management.

XML Encryption & XML Digital Signatures

The metadata associated with culturally sensitive Indigenous resources is often sensitive itself. It is important that the integrity and confidentiality of this data is guaranteed. It is often the case

that for a single resource there may be both confidential and non-confidential metadata.

XML Encryption and XML Digital Signatures [11][12] provide a mechanism for protecting these resources. They allow encryption of selected pieces of metadata, as well as the actual resource.

2.1.4 Multimedia Presentation

SMIL

The Synchronized Multimedia Integration Language [13] enables simple authoring of interactive audiovisual presentations. Using a simple markup language it is possible to generate multimedia presentations consisting of text, images, audio and video. It allows for the synchronisation of the different multimedia content being presented. Using SMIL it is possible to dynamically generate multimedia presentations of Indigenous resources.

2.2 Unique Requirements of Indigenous Knowledge

Investigation into the laws and constraints governing Indigenous intellectual property is important in devising a model and tools for their protection. A number of papers [4][14][15] have been written that attempt to explain both the unique intellectual property requirements of Indigenous knowledge, and the laws and protocols that exist in Indigenous communities.

Githaiga [14] and Janke [15] have identified how Australian copyright laws fail to fully accommodate and protect Indigenous knowledge. Similar inadequacies exist for many other countries:

Ownership and Authoring Western copyright laws are primarily based on the identification of an individual owner or creator, focusing on their entitlements to the economic awards for their creations. The Indigenous view is based on collective, rather than individual ownership. In many cases no single individual can be identified as the owner or author as the work is based on cultural traditions which have been handed down over generations.

Originality Many Indigenous works are based on stories, traditions, symbols, designs or techniques that have been replicated over time. The notion of originality required by modern copyright laws that reflects the “distinctive individual creative style” of the author often does not exist for Indigenous works.

Material Form Indigenous knowledge often exists in oral or visual form, rather than expressed in written documents. Hence the condition of fixation in material form excludes many Indigenous works from protection under modern copyright laws.

Duration Modern copyright laws generally limit the duration of copyright protection to the creator's life plus a fixed period (50 years in Australia). After this period the resources are made freely available. This fixed period is inadequate for Indigenous resources for which community ownership exists in perpetuity.

Rights of Derivative Works Modern copyright laws consider derivative works as original creations and therefore require that they have their own copyright. If the traditional owners are not the legal owners of an Indigenous resource then they have no control over how others use, adapt or reproduce their work.

Customary Law The most significant inadequacy of modern copyright laws are that they do not support the unique rules and protocols governing access to Indigenous knowledge. These traditional rules are typically decided by senior members of the tribal group and are often very complex. There are a number of common factors amongst Aboriginal communities determining access to resources. These include:

- the person's membership to a particular tribe, clan, band or family;
- the person's status within that tribal group: e.g. whether they are a child or elder;
- the person's role within the tribal group: e.g. whether they are artist, musician, midwife.
- the person's gender;
- the person's relationship to the objects depicted in the resource: e.g. in some Indigenous communities it is inappropriate for male members to view images of their mother-in-law;
- the death of people depicted in the resource. Some Indigenous cultures enforce the law of "sorry business" by which it is inappropriate to view images of people who have died for a fixed "mourning" period;
- the context in which the resource is to be used: e.g. it may be acceptable for anyone to view and copy a particular traditional work of art, but it would not be considered appropriate to use it in commercial advertising.

Dynamic Another significant problem is that Indigenous laws and protocols are dynamic. Typically the laws are decided by a small group of elders in the community and as a result they change over time. This is very different to western copyright law which is typically static. Modern technologies are not designed to accommodate this volatility.

Heterogeneous Not only do Indigenous laws differ over time they can differ vastly between communities. Although many commonalities exist it is impossible to state a standard set of protocols and laws.

Secret Laws In some cases the laws and protocols that govern access to certain resources are secret themselves and known only to a select few. Having the rules that govern access control confidential themselves creates a security paradigm not generally considered by modern technologies.

Tribal Care Laws Staff at the National Museum of the American Indian (NMAI) in Washington also expressed the need to be able to document traditional care information associated with culturally significant physical artefacts in their custody. Indigenous tribal representatives frequently requested traditional care constraints for the storage and handling of culturally significant or sensitive objects. John Pepper Henry, the Repatriation Program Manager at the NMAI provided this list of common constraints:

- directional orientation;
- segregation from other objects or other tribes' materials;
- storage on higher shelves, use of wooden shelving;
- cover from view with cloth or muslin;
- ceremonial feeding/dusting with corn pollen and/or other materials;
- cleansing with smoke (cedar, sage, tobacco, sweetgrass etc.);
- freezing restriction;
- low oxygen restriction (no CO2 bubble or covering with plastic or other non permeable materials);
- sacred/ceremonial bundle dis-assemblage restriction (bundle to remain intact and unopened);

- ceremonial pipe assemblage restriction (bowl and stem should be housed disconnected).

From these comparisons it is clear that modern intellectual property laws differ vastly from Indigenous laws and protocol. The technologies described in this chapter are all based on modern intellectual property laws and in many cases do not accommodate the requirements of Indigenous knowledge.

2.3 Rights Extensions

The use of modern technology in the protection of Indigenous knowledge was considered in the paper *Rights Markup Extensions for the Protection of Indigenous Knowledge* by Hunter [16]. This paper is one of the few pieces of literature that looks at the inadequacies of modern technologies in dealing with Indigenous knowledge. It focuses on current international standards for rights expression (XrML, ODRL and MPEG-21) and explores possible extensions to these technologies to accommodate Indigenous knowledge. This work provides the theorectical background for this thesis and is a starting point for further research into both the limitations of modern technologies and how they may be overcome.

Additional work by Hunter and James in *The Application of an Event-Aware Metadata Model to an Online Oral History Project* [17] provided information into a possible design and architecture for the proposed software tools.

Chapter 3

Software Requirements Specification

As identified in the previous chapter the requirements pertaining to the use and management of Indigenous knowledge are complex and unique. Consequently software that deals with Indigenous knowledge comes with its own set of unique requirements. This chapter investigates the software requirements for developing a system for the management of Indigenous resources.

3.1 Intended Usage / Environment

As a system for managing Indigenous knowledge is not targeted at a single homogenous group of users, its intended environment and usage is difficult to define. However it is possible to describe an envisaged usage and state how this may differ depending on its particular application.

3.1.1 Envisaged Usage

Figure 3.1 shows the envisaged process of how software tools may be used:

1. A particular community or cultural institution has a collection of resources they would like to have digitally or physically repatriated. Some communities may want only the digital representation of the physical object, others may want both the physical and digital copy, while others may be happy to have both kept by the museum or cultural institution;

2. The resources are digitised and any cataloguing or indexing information entered. This may be done from existing catalogue information stored in museums, libraries or other cultural institutions.

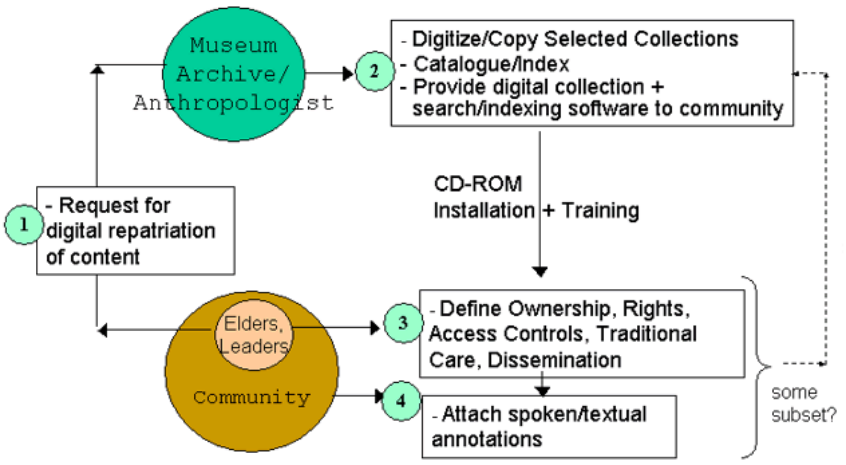

Figure 3.1: Envisaged use of software tools.

The digitised resources, metadata and software tools are then made available to the community concerned;

3. The elders, leaders or selected representatives of the community can then define their particular rights requirements and create user profiles governing who is able to access which resources and under what conditions.

   They can also add or alter existing catalogue information currently stored in the system;

4. Authorised users may then attach spoken or written annotation to resources of interest;

5. If the community wishes to make all or some of the content they have entered available to other cultural institutions or the general public they may do so.

3.1.2 Users and Roles

A number of different users may access the systems, some examples are provided below:

- a librarian or museum curator may digitise a collection and enter cataloguing (Dublin Core) metadata describing the resource;
- an anthropologist may add written annotations describing the resources;
- a system administrator may create a number of user accounts;

- a traditional owner may define rights constraints pertaining to the collection;
- an Indigenous user may add spoken annotations describing the collection;
- a member of the general public may search the collection and view items (without rights restrictions) as a multimedia presentation.

As a result three user types have been created with following roles:

Admin users have open access to the system. They may enter all forms of metadata, as well as create user profiles. Admin users may be system administrators, library or museum staff or the traditional owners of the collection;

Owner users are able to define the rights constraints pertaining to particular resources as well as enter annotations and tribal care metadata. Owner users will typically be the traditional owners or caretakers of the resources;

Users are simply allowed to add annotations and view those resources they have permission to access.

3.1.3 Scope of Usage

Due to the diversity of Indigenous communities and the variations in their particular requirements the process described above may be altered to suit a particular groups needs. Some communities may be content with the museum or cultural centre keeping both the physical and digital resources, as well as all the software. Others may want to be the sole holders of all three.

In some cases the system will be required to run on a single stand-alone machine. This machine could be a laptop that may be used by a single person travelling to remote Indigenous communities. In contrast to this, a large infrastructure may be in place where the system may be distributed across a number of different domains, with the metadata, resources, and clients all at separated location.

From these examples it is clear that the intended environment could differ significantly depending on its particular application. As a result it should be designed to accommodate a wide range of different possible scenarios of usage.

3.2 Functional Requirements

Reviewing the objectives presented in Section 1.1 the functional requirements can be split into two separate categories described below.

3.2.1 Metadata Input and System Administration (Back-end)

An application is needed where authorised users can log on and based on their profile enter metadata on a particular resource, as well as administer user accounts and rights constraints. More specifically the system needs to allow users to:

- open different multimedia resources (video, audio, images, text) and view them;
- enter descriptive (Dublin Core) metadata about a resource;
- record both written and spoken annotations and attach them to a resource, as well as allowing users to listen/view annotations for a particular resource;
- restrict access to sacred or secret resources by specifying constraints that must match the user profile of the person wanting to access the resource;
- specify customary rules and protocols needed to govern access control to particular resources;
- create user accounts and profiles that describe users, so that the system may control the content they are presented with;
- perform keyword searches on the metadata to find resources of interest.

3.2.2 Search, Browse & Retrieval Interface (Front-end)

To provide general access to the resources and metadata describing them a search and retrieval interface is needed. This system should:

- allow users to log on and perform keyword searches;
- filter results based on the rights restrictions of resources and the user's profiles;
- allow users to view or listen to the list of annotations associated with a resource;

- allow user's to add resources of particular interest to their "shopping cart" or custom collection so that they may later easily access them;
- allow users to select resources they wish to be presented within a dynamically generated multimedia presentation.

3.3 Overall System Requirements

The nature of Indigenous knowledge introduces a number of specific software requirements that are important to the success of the proposed software tools:

Robustness The software should be able to stand up to the rigours of unexpected input by users with low computer literacy skills.

Low Cost In order to make the software open source and available to Indigenous and grassroots communities it is essential that it is built as inexpensively as possible.

Interoperability The software should be built on the international standards described in Section 2.1. This ensures maximum interoperability between different systems and allows for easy future development or expansion. It should run on Unix, Macintosh and Windows operating systems.

Portability The software should be designed to operate on a wide range of different platforms and operating systems. The software should be ported with as little effort as possible.

Scalability As identified, the scale to which the system may be used can vary greatly - from a single stand-alone machine to a wide-area distributed network. As a result the software should scale efficiently as the size of the system grows.

Flexibility One of the most important attributes of the software is that it is customisable. Due to the complex nature of Indigenous knowledge it is essential that users are able to adapt the underlying data model (or schema) to their particular needs.

3.4 Software & Hardware Interface Requirements

This section defines the software's requirements in terms of other software and hardware systems that it must interface with. Due to the nature of the software's system attributes (interoperable, portable, etc.) these may vary greatly.

3.4.1 Software

The software interface requirements are mainly determined by the Overall System Requirements in the previous section (3.3):

- To ensure the system is reliable, third party software should be tried and tested;
- Interoperability and portability requirements mean the software that the system uses needs to be platform independent and based on international standards;
- The use of only open source software and platforms would be most advantageous in ensuring minimal costs;
- The scalability of third party software should be taken into account;
- The system should use software that is flexible, allowing users to alter the underlying data model. User interfaces should reflect the changes to the data model.

3.4.2 Hardware

Due to platform independence and interoperability requirements of the software, there are few specific hardware requirements. The main hardware requirement is ensuring the system has enough hardware resources to support the software. The hardware resources required are highly dependent on how the system is implemented in its target environment. Some key hardware resources are outlined below:

Digitisation Equipment is not specifically to be used by the proposed software tools. However digital scanners and cameras need to support data formats that are readable by the software, and are of acceptable size and quality.

Processing and Memory resources should be sufficient to run the software. As a stand-alone installation this may simply require enough CPU power and memory to run the software

for a single user. In contrast if the system is distributed over a network CPU and memory resources should be sufficient to support many concurrent users from different locations.

Secondary Storage is largely based on the size of the collection that has been digitised. Multimedia resources (especially high resolution video) requires large storage capacity. Sufficient secondary storage systems needs to be available. The resources may also be made available on different storage mediums, e.g. CDROM, DVD, ZIP drive.

Network requirements depend on how the system is to be used. Stand-alone, single user systems require no network access. However if the system is distributed over a wide-area network with many concurrent users then high bandwidth reliable communications are needed. The system design should also take into account the effect of low bandwidth or unreliable communications and how this may effect the transport of larger audio or video multimedia resources.

3.5 User Interface

The user interface should reflect the different user roles identified in the system:

- Many potential users will have low computer literacy skills, thus the system must have simple, intuitive user interfaces. The use of simple, well known interfaces such as web-based interfaces should be considered;
- The user interface should differ depending on the user type (Admin, Owner, User). Users should only be presented only with the functions they are allowed to perform;
- The interface should take into account the range of different multimedia types that may exist and how they are best presented.

3.6 Security Requirements

Due to the sensitive nature of possible collections it is important that the system has the correct mechanisms in place to protect resources from misuse. Collections containing highly sensitive material may be required to run on a stand-alone machine or intranet rather than on public internets. If the system is to be used over an internet, sufficient communication security mechanisms

need to in place. Public/private key encryption and digital signatures may be used to to ensure integrity and confidentiality of communications.

3.6.1 Trust Requirements

The security functions described in Section 3.2.1 regarding adding and managing user profiles creates an important *trust* requirement. The security of the system is dependant on the accuracy of the information in user profiles. A trusted person is required to both enter profiles correctly and validate the person for whom the profile is being created. This *trust* requirements cannot be implemented in software, but it is important that it is upheld in the systems intended environment.

Chapter 4

Data Model & System Architecture

The previous chapter identifed the software requirements. In this chapter a high level view of the software is given.

A flexible underlying data model, showing how metadata is represented in the system is described; a high level system architecture, workflow and user interaction is presented; and a explanation of how the presented data model is mapped to a relational database is given.

4.1 Metadata Model

Figure 4.1 shows the three different metadata types of the system. From this figure it is clear that the resource represented can be a purely digital resource (e.g. a photograph taken with a digital camera), or a digital representation of a physical object (e.g. a photograph of a sacred artefact). This needs to be considered when designing a metadata model to represent the resource. The three metadata categories shown in the figure are described below:

Dublin Core + CIDOC CRM (Descriptive) Metadata contains metadata that simply describes the object. This is based on the 15 fields of the Dublin Core standard. Current work involves also incorporating the CIDOC CRM standard to allow detailed descriptions of physical objects.

Rights / Tribal Care Metadata contains metadata describing how the resource may be used and under what conditions. XrML syntax is used to specify rights constraints. As the resource may represent both a physical or purely digital object, the rights metadata needs to be divided into one of three categories:

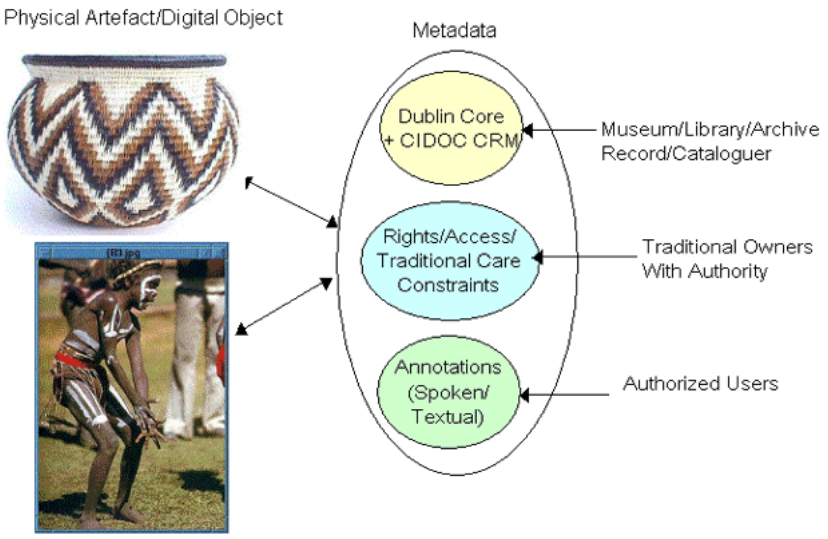

Figure 4.1: Metadata categories

Viewing / Listening represents the metadata describing the constraints regarding access to view or listen to the resource. This metadata is what the software uses to govern access control.

Copying represents the metadata describing the constraints regarding the copying of the resource. If the resource is a digital representation of a physical object this category may be used to represent who may take photographs / create likenesses of the physical object.

Handling represents the metadata describing the constraints regarding the handling of the resource. This may represent constraints on who may touch or move sacred artefacts or other physical objects.

Two additional types of rights metadata are also added that allow tradition owners additional information regarding the use of a resource:

Request represents any requests traditional owners may have regarding the use of a resource. “Only to be reproduced for non-profit purposes” may be an example of a type of request.

Warning represents any warnings that should be displayed to people wanting to view the resource. If the resource contains culturally sensitive material traditional owners may

include a warning.

Three categories: *Orientation*, *Storage* and *Preservation* are provided to represent Tribal Care constraints. These are described in Chapter 5.

Annotative Metadata contains the annotations added to a resource. The metadata captures who added the annotation, when it was added and the text / audio content of the annotation.

Due to the complex and dynamic nature of the laws and protocols governing the use of Indigenous knowledge, the rights metadata needs to be flexible and easily customised by traditional owners. Using XrML syntax to describe rights constraints and providing the five rights categories described above allows the rights metadata to be flexible and easily customised.

4.2 System Architecture

Based on the requirements identified in the previous chapter a system has been developed which consist of three major components:

Metadata Editor / Generator Application A Java application that allows authorised users to log in and based on their profile enter the three types of metadata described in the previous section. The application also provides system administrative facilities for managing user profiles and allows users to customise the rights metadata model to suit their particular needs.

Java was chosen for its portable features and good API support for: media with the Java Media Framework (JMF [18]); XML with the JDOM API [19]; and database access with JDBC [20]. The Metadata Editor / Generator Application is described in detail in Chapter 5.

Database The metadata generated by the Java application is stored in relational tables implemented as a MySQL [21] database and accessed using the Java Database Connectivity (JDBC [20]) API. MySQL was chosen because it was open source and has good library support for Java. The specifics of the database structure is explained in Section 4.3.

Search, Browse & Retrieval Interface A number of dynamically generated webpages have been created to facilitate searching and presentation of resources stored in the MySQL database. These pages are implemented in Java Server Pages (JSP [22]). JSP was chosen for similar

reasons as Java in the Metadata Editor / Generator Application. The Search, Browse & Retrieval Interface is described in detail in Chapter 6.

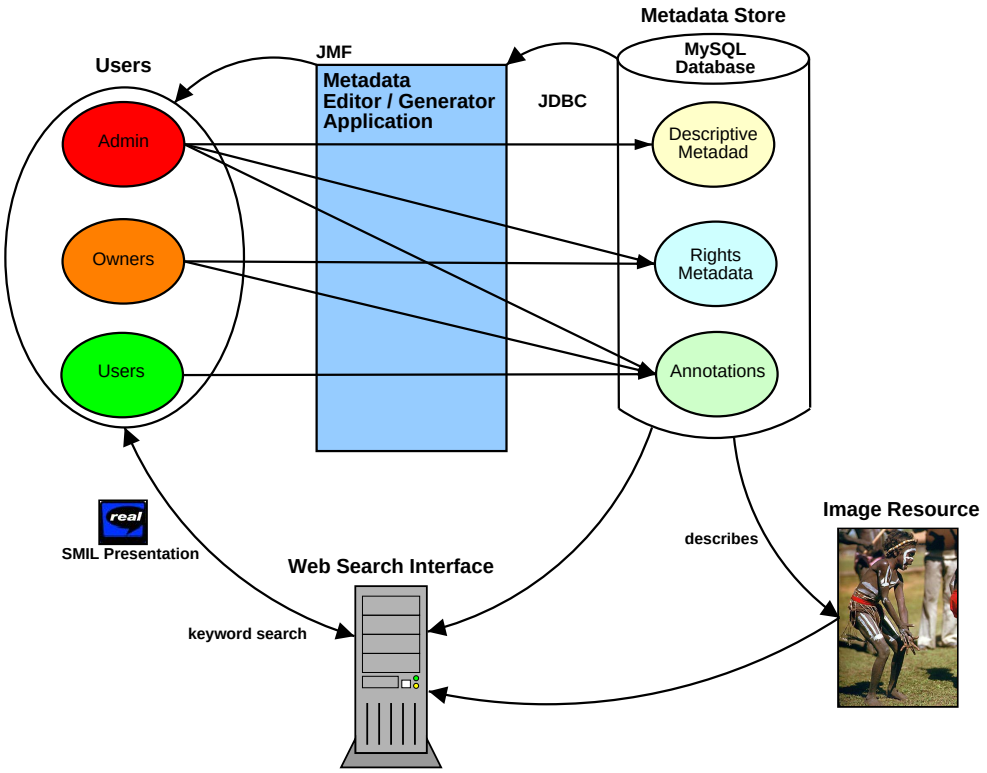

Figure 4.2: User Interaction

Figure 4.2 and Figure 4.3 shows the user interaction and high level architecture of the system:

- The resources are digitised using different capture devices and are stored as multimedia files on secondary storage;
- Using the Metadata Editor / Generator admin users create user profiles and edit the rights schema to suite the requirements of the particular collection;
- Users log on, open the multimedia files, and input descriptive, rights, and annotative metadata. The user type (admin, owner, user) determines the type of metadata the user is allowed to enter;
- Web users logon and perform searches on the metadata stored in the database. Results are filtered according to the user's profile;

- The results may be viewed as a dynamically generated SMIL presentation.

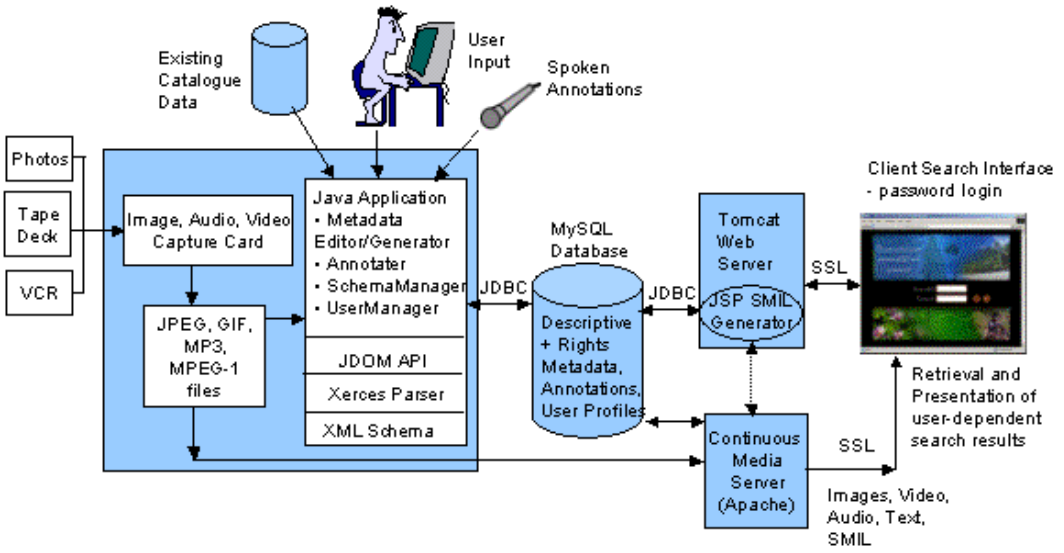

Figure 4.3: System Architecture

4.3 Database

An important design decision was whether to store metadata purely as XML documents or to use a relational database system. There are advantages and disadvantages to both forms of storage. On review, a decision to use a relational database system was made because:

- searching data in database tables is easier and far more efficient than searching across XML documents. When considering large or multiple collections there is considerable overhead in performing keyword searches. The use of SQL on relational tables allows for fast and complex searches which are easy to create. Similar query languages do exist for XML (such as XPath and XQuery), but take considerably longer because they have to open each XML document and navigate the Document Object Model. Other performance improvement techniques such as indexing and clustering are also only available to relational database systems;

- in addition to the simplicity and speed of searches, databases allow simple data input and manipulation. The JDBC API allows Java applications (or JSP servlets) to insert data and perform queries by simple SQL statements. These technologies are well documented and are widely used;
- if there is existing metadata describing a digital collection in another system it will most likely also be stored in a relational database. Importing or interfacing with these systems is made simpler if a relational database is used;
- the requirements for robustness stated in the previous chapter is better met by the use of a database. A central data store with backup and recovery facilities is more reliable than separate XML files that could be stored in different locations;
- managing access to a single database can easily be done using DBMS user accounts and privileges. Databases provide a far more secure system when considering the sacred and secret nature of the system.

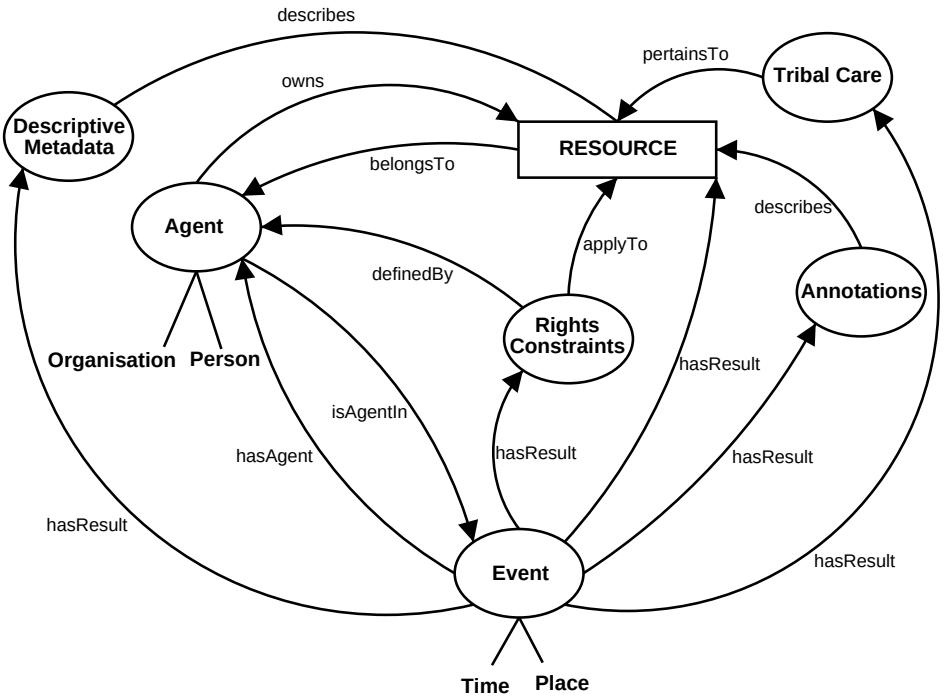

Figure 4.4: Abstract representation of the database structure

Figure 4.4 gives a high level view of the database structure. The event-based behaviour is explained in the next section. For the details of the database structure the SQL create scripts have been provided as Appendix D.

The disadvantage of storing the metadata in relational tables is that the database schema, once created, is static. To make the software truly flexible and customisable users need to be able to customise the underlying data structure.

This could be overcome by providing a small tool which administrator users can use to customise the structure of the database. Although the core structure would be explicitly set, the tool may provide a simple interface to add custom fields. From the user's selection SQL scripts could be produced and may be run to create the database tables. This could be done only once when the system is installed and before users begin entering metadata.

4.3.1 Event-based System

In light of the security requirements and given that the systems has different types of users, it was determined important that the system should record certain events when they are performed. These include when:

- users login and logout of the system;
- users enter descriptive metadata;
- users add annotations;
- users add / change rights restrictions;
- users add / change tribal care recommendations;
- administrators create or edit user profiles;
- administrators customise the rights schema;

For this purpose an `EVENTS` table was created that documents:

- the type of event: *descriptive metadata*, *rights*, *tribal care*, *annotation*, *login*, *schema*, *user profile*.
- the date and time the event occurred;

- who performed the event;
- the resource on which the event was performed on: e.g. the ID of the resource who's rights constraints have been changed;
- the result or output of the event: e.g. in the case of a rights event a pointer to an entry in the `RIGHTS` table that contains the new rights restriction information.

Figure 4.5 gives an example of the contents of the database tables after a user *bkoopman* has entered a *rights_event* to a resource *resource1* with rights restrictions given by rights ID *3*.

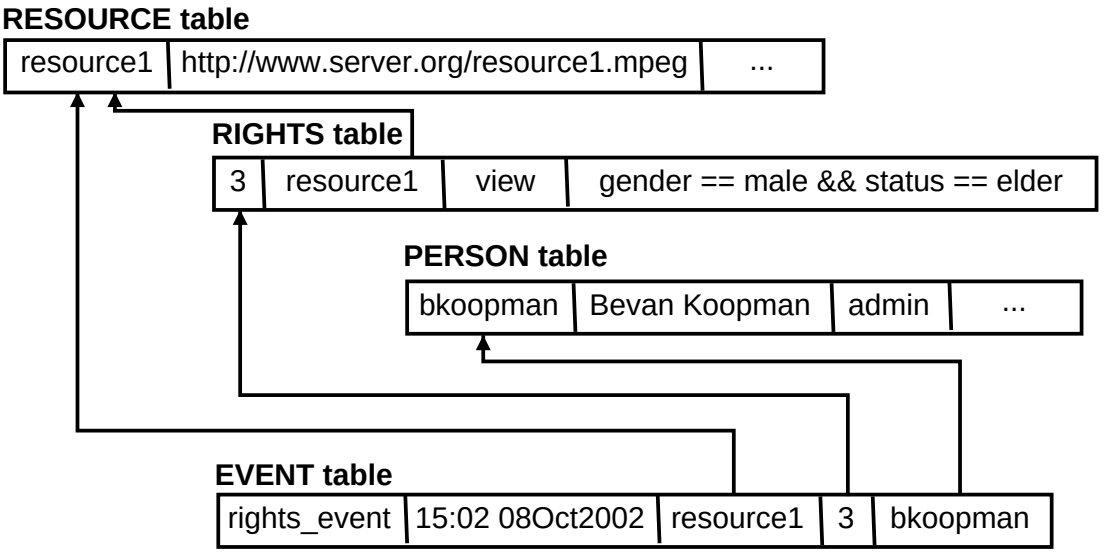

Figure 4.5: Example of a *rights* event entry

Chapter 5

Metadata Editor/Generator Software

The Metadata Editor/Generator is a Java application that enables users to enter the three different types of metadata: descriptive, rights and annotative. It uses JDBC (Java Database Connectivity) [20] to access the metadata stored in the MySQL [21] database. It provides a search facility to find specific resources and has a resource viewer so that users may view / listen to resources. The administrative component of the application allows authorised users to manage user profiles and customise the schema that defines the rights metadata.

5.1 Architecture

Figure 5.1 gives an overview of the different components of the Metadata Editor / Generator application, how they relate to each other, and the respective technologies / APIs used.

Flexibility and portability have been identified as important attributes of the system. To facilitate this an XML configuration file is provided that is read by the application. The XML file contains:

- a JDBC URL that refers to a MySQL database server and specific MySQL database where the metadata is (to be) stored;
- a username and password to use to connect to the MySQL database;
- the path to a directory where a collection of resources (multimedia files) are stored;
- the URL to a webserver where the resources can be retrieved. This is needed by the Search, Browse and Retrieval interface described in the next chapter;

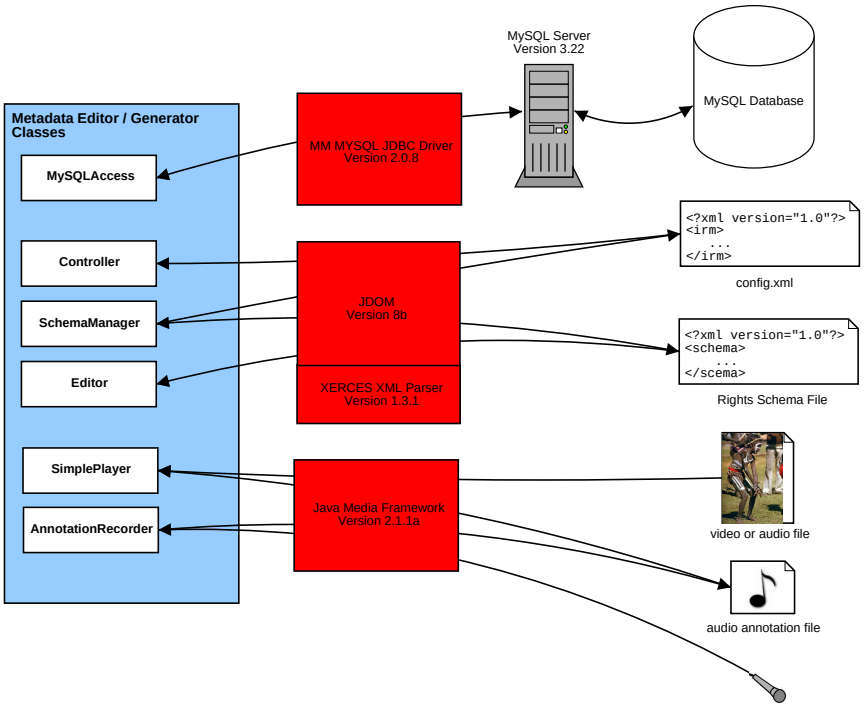

Figure 5.1: Metadata Editor / Generator Application.

- the current XML schema file that the system is using. This is set by the SchemaManager.

An example of the configuration file is provided as Appendix A.

Figure 5.2 shows a hierarchy of the different components of the application. The remainder of this chapter involves more detailed descriptions of the main components of the Metadata Editor / Generator application.

5.2 User Profiles

Every user has an associated user profile which contains the information about the particular user. Currently the information represented includes:

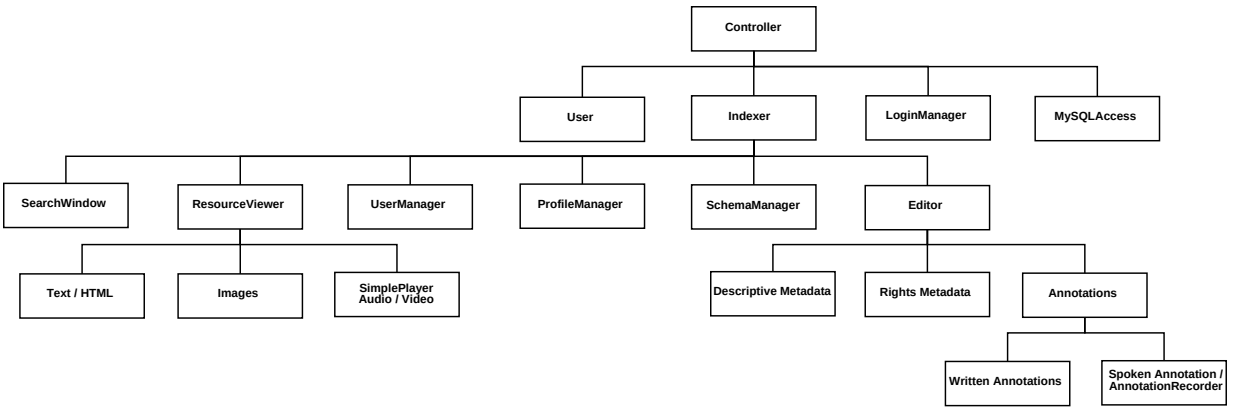

Figure 5.2: Metadata Editor/Generator Hierarchy

- login id;
- name;
- user type: admin, owner or user;
- password;
- organisation;
- tribe;
- clan;
- band;
- family;
- birth and/or death date;
- gender;
- status: e.g. elder, child;
- role: e.g. dancer, artist;
- mother;
- father;
- spouse.

The facilities available to a particular user depend on their type.

5.2.1 UserManager

All user profile information must be entered by an Admin user when the user's profile is created. Only Admin users have access to the UserManager.

Rather than creating individual accounts for each user, Admin users may create a profile for a group of users. The user id and password can then be given to the respective group and all members may use the account. This allows administrators to create profiles for a single homogenous group. Figure 5.3 shows the UserManager for a elder member of the Nunukal tribe.

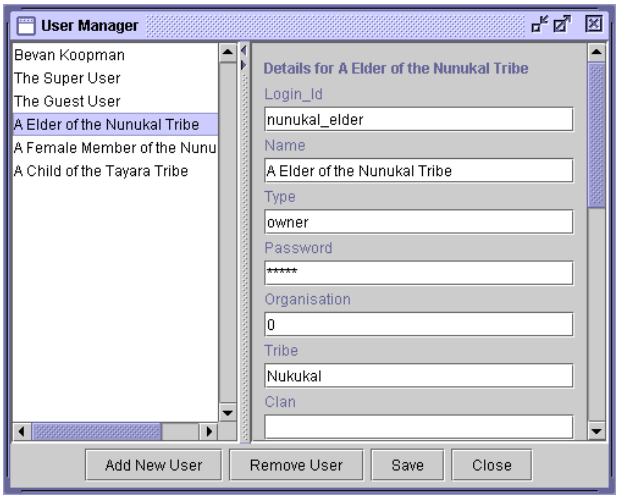

Figure 5.3: UserManager

5.2.2 ProfileManager

The ProfileManager was created so that users may manage their own profile. At present all they are able to do is change their password. It is important not to allow users to change other attributes as they could alter their profile to gain access to a resource they should not view.

The ProfileManager was also created with the idea that future developments may allow users to customise the way they interact with the system. Users may differ vastly, and interact with the system in different ways. A librarian, for example may only be interested in entering descriptive metadata, whereas an Indigenous user from a remote community may be wanting to attach spoken annotations in their own language. The ProfileManager could allow these differing users to customise the software to their particular needs. For example, based on information entered in

the ProfileManager users could be presented with different graphical user interfaces: Indigenous users in a remote community with low computer literacy skills may be presented with a simpler interface, with advanced features not visible. In contrast to this, advanced users may wish to have all user messages turned on so that they may know exactly what the systems is doing.

The ProfileManager helps to ensure flexibility of the software, which is essential when dealing with the complexities of Indigenous knowledge management.

5.3 SchemaManager

The main component of the system that facilitates customisation of the software is the SchemaManager. It allows Admin and Owner users to specify the rights restrictions used by the Editor. Its features include:

- allowing users to enable and disable certain constraints (Figure 5.4). If a particular constraint is enabled then it will appear in the *Rights* tab of the Editor, thus allow restrictions based on this constraint.

- allowing users to specify new rights constraints (Figure 5.5) pertinent to their particular needs. Users may enter the *name*, *type* and *vocabulary* of new rights options. The *type* may be one of three options:

  value the new rights item may have any value. e.g. Tribal Name may be any text string;

  disjoint the new rights item can only have a single value from the specified vocabulary. e.g. Gender may be male OR female;

  intersect the new rights item can have any value(s) from the specified vocabulary. e.g. Role may Dancer AND Artist.

- allowing users to save this information to a XML file, an example of which is shown in Appendix B. This XML Schema file is used by the Editor to generate its interface with the correct rights item enabled.

- allowing uses to load a particular XML schema file they wish to use. The SchemaManager will then read the schema file (using the JDOM API [19]) and use the rights constraint specified in the loaded file.

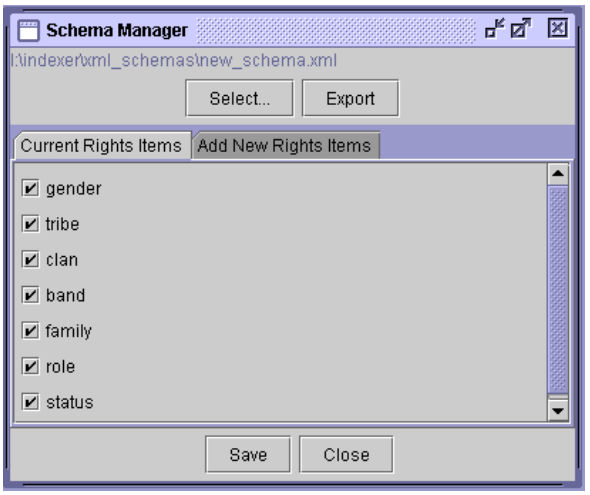

Figure 5.4: SchemaManager: enabling rights restrictions

As we have identified, the requirements of Indigenous communities differs not only from western communities but also between different Indigenous communities themselves. The Schema-Manager means that users can alter the software to suit their needs, as well as allowing users to simply move between different configurations by simply loading new schemas into the system.

5.4 Descriptive Metadata

The descriptive metadata is comprised of the 15 fields of the Dublin Core [5] and an additional field *Resource ID* used to give the particular resource a unique identifier. This metadata may be entered by users of the system, or could come from existing catalogue data. The *Identifier* field of the Dublin Core field is used to store a URL to the particular resource. This URL is needed by the Search, Browse & Retrieval Interface described in Chapter 6.

Current work involves extending this section to include relevant parts of the CIDOC CRC described in Section 2.1.1. This will allow for detailed descriptions of physical objects that are represented in digital form, such as photographs of valuable artefacts.

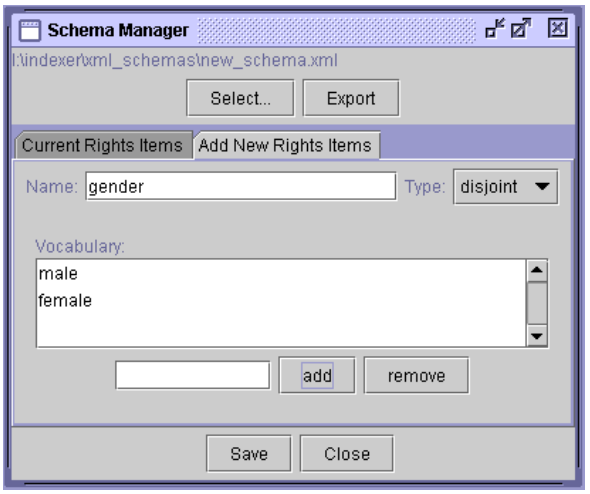

Figure 5.5: SchemaManager: adding new rights restrictions

5.5 Rights Metadata

The user interface that allows users to enter rights metadata is generated from the XML schema created / loaded by the SchemaManager. For each `<rights>` item (See Appendix B) a corresponding label and checkbox are created allowing the user to enable the rights constraint. A textfield, drop-down box or list are created depending on the rights item's `type`. These widgets are populated with the item's vocabulary elements. An example of the RightsEditor is shown in Figure 5.7.

The rights metadata entered using this interface can be one of three types:

Viewing/Listening is rights metadata associated with who can view, or in the case of audio/video resources listen to a particular resource. The rights checking of both the Metadata Editor/Generator and the Search & Retrieval Interface is done on Viewing/Listening rights metadata;

Copying is rights metadata associated with who is able to makes copies of the resource. In the case of physical objects this constraint may be used to represent who is able to photograph/sketch pictures of the resource;

Handling is rights metadata associated with who is able to handle the physical object. This category could be used to govern who is able to keep a copy of the digital object.

The *Ritual* tab of the rights editor (Figure 5.8) provides support for:

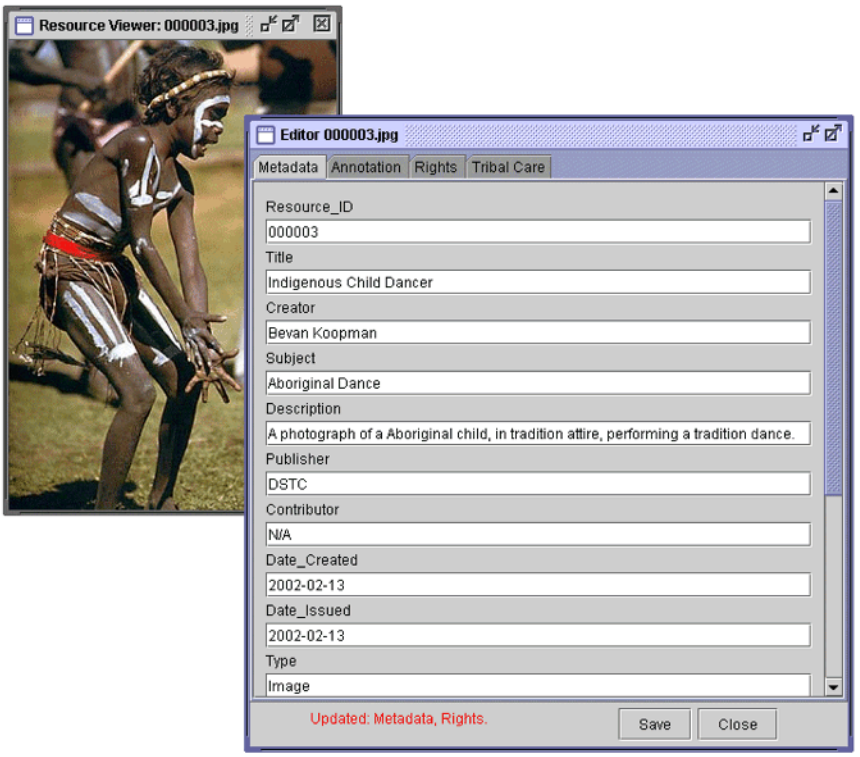

Figure 5.6: Descriptive (Dublin Core) metadata for a image.

- restricting access to the particular resource for a period of time. This is in accordance with the law of “sorry business” that some Indigenous communities enforce. The law states that no one should view / listen to resources that depict someone who has died;
- attaching a textual warning to the resource that will displayed before the resource is viewed.

Another text field is also added to the RightsEditor to allow owners or administrators to add any specific contextual constraints that may be associated with the resource (Figure 5.9). These constraints would be displayed to users before viewing the resource.

5.6 Tribal Care Metadata

In response to requests from both museum staff and representatives of Indigenous communities, we have also added support for the specification of Tribal Care recommendations (as identified by Dunman [23]) by the original owners of resources. The Tribal Care metadata tool enables

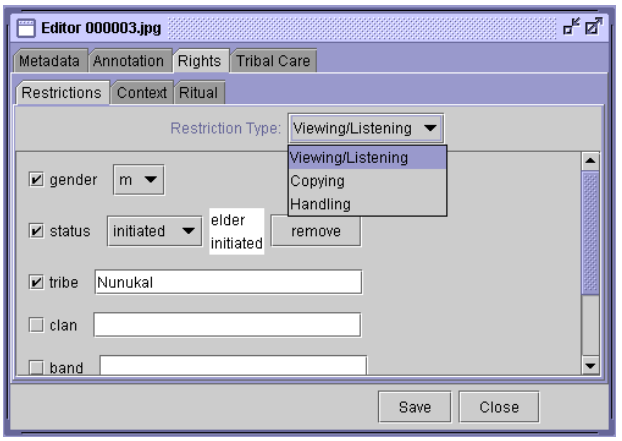

Figure 5.7: RightsEditor: restrictions on gender, status and role enabled

Figure 5.8: RightsEditor: warning and “sorry business” restrictions

Indigenous communities to define the ways in which culturally significant or sensitive artefacts should be treated, ways which are acceptable to the cultures represented.

Currently the application provides support the three different categories of Tribal Care:

Spatial Orientation allows users to specify how the physical object should be orientated, whether it needs to be kept facing a certain direction or placed in a certain part of the room.

Storage allows users to specify specific requirements on how the physical object is to be stored. If it needs to be stored on or in any other specific material or if it should be segregated from other types of objects.

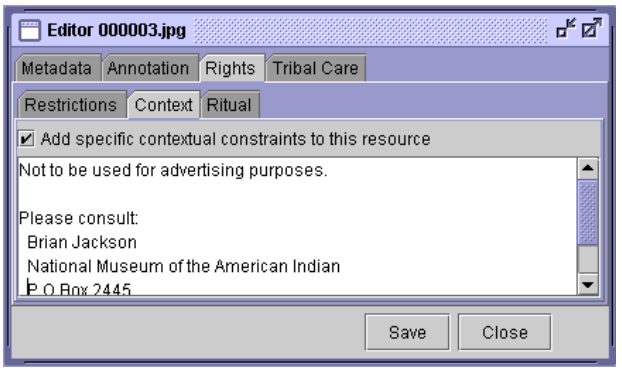

Figure 5.9: RightsEditor: specific contextual constraints

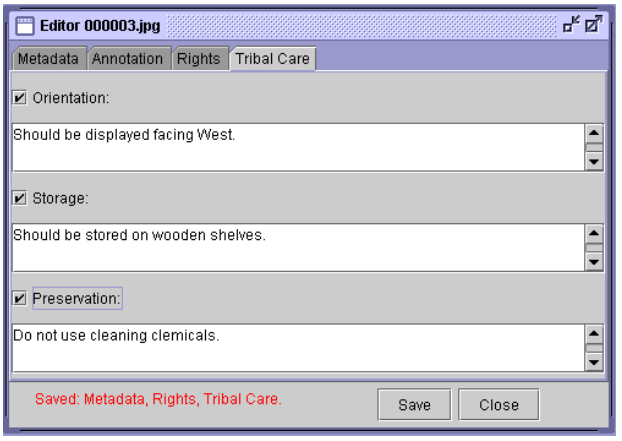

Figure 5.10: Examples of Tribal Care recommendations.

Preservation allows users to specify specific preservation and conservation requirements. These may be physical (for example to be kept in a dry environment) or spiritual (for example to be used one a year in a particular tribal ceremony).

Figure 5.10 shows the Tribal Care editor with some examples of types of recommendations to ensure the respectful tribal care of physical artefacts.

While the actual implementation of traditional care specifications may be difficult, many museums are attempting to integrate Indigenous beliefs and requests into museum practices and the software will facilitate this process.

5.7 Annotations

As mentioned in the functional requirements for metadata input (Section 3.2.1), allowing users to add annotations to a particular resource is an important feature of a system for managing Indigenous knowledge. Annotations become resources themselves as they provide valuable additional metadata about the resource. In Indigenous communities much of the knowledge is in the minds of the people, where the prefered method of communication is through spoken stories or songs that have been passed down through generations. Annotation tools provide a way of capturing these oral traditions, making them available to others, and preserving them for future generations.

As a result the Metadata/Editor Generator application provides the following facilities:

- allows users to add a written (textual) annotation to a particular resource;
- using the AnnotationRecorder allow users to attach a spoken annotation to the resource through a microphone. Figure 5.11 shows a written annotation and Figure 5.12 shows a recording of a spoken annotation in progress;
- allow users to listen / view annotations that have been added to the resource. Information on who added the resource, and when it was added is also displayed;
- allows Admin users to delete unwanted annotations.

The AnnotationRecorder, and SimplePlayer (for playback of annotations) are both implemented using the Java Media Framework (JMF) [18], allowing them to be platform independent.

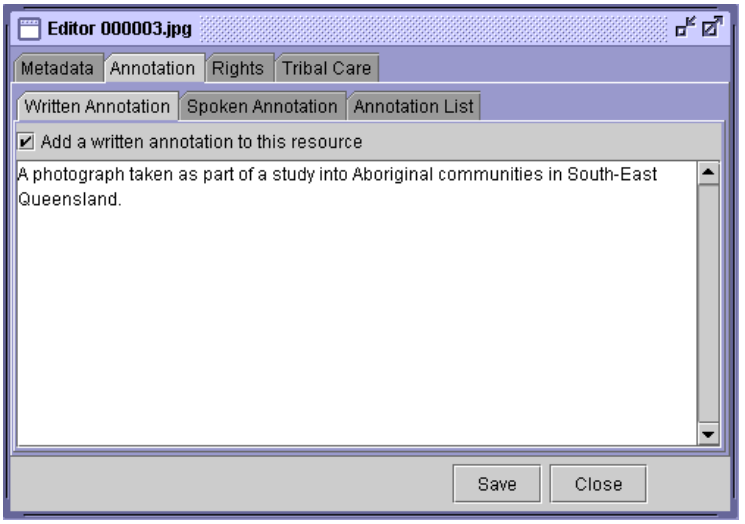

Figure 5.11: AnnotationRecorder: written annotation.

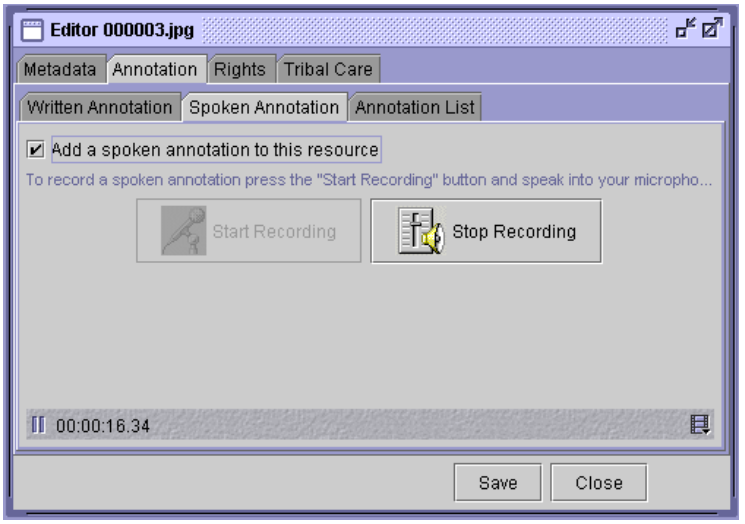

Figure 5.12: AnnotationRecorder: currently recording a spoken annotation.

Chapter 6

Search, Browse and Retrieval Interface

To allow access to collections of media items that have been indexed using the Metadata Editor/Generator described in the previous chapter a web-based interface has been developed. A number of dynamically generated webpages are created using Java Server Pages (JSP [22]). These pages allow users to logon, and based on their profile (as created by the UserManager of Section 5.2.1) perform searches on the metadata contained in the system's MySQL database. The results returned are filtered to include only those the user has permission to access. User's may then add media items of interest to their collection and be presented with dynamically generated multimedia (SMIL [13]) presentations.

6.1 Architecture

Figure 6.1 shows the architecture of the Search, Browse and Retrieval Interface:

- the client's web browser submits requests via an internet or intranet. Communications are encrypted using Secure Socket Layers (SSL [24]);
- server side processing is done by Java Server Pages (JSP [22]) on a Jakarta Tomcat webserver.
- the Tomcat server uses JDBC to connect to the MySQL database containing the system's metadata.
- because SMIL is an XML based markup language it can be navigated and manipulated using the JDOM API [19]. The JSP pages running on the Tomcat server can therefore read

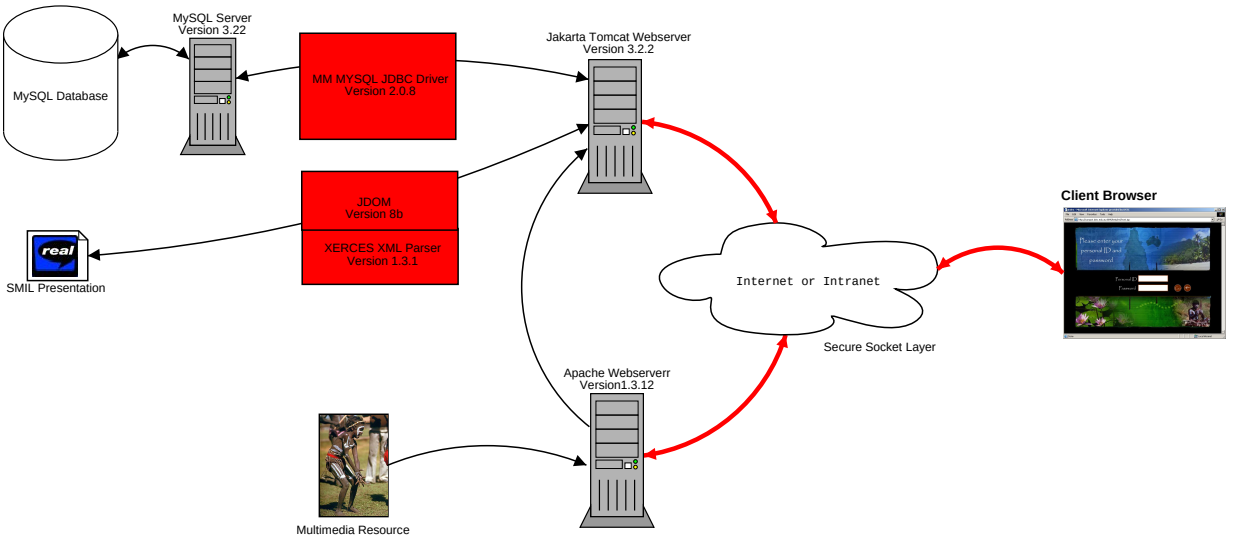

Figure 6.1: Architecture of the Search, Browse & Retrieval Interface

SMIL templates and manipulate the Document Object Model to create custom presentations.

- the additional Apache webserver pictured in Figure 6.1 provides access to the actual multimedia resources. It is quite possible for the Tomcat webserver to perform server-side processing and provide HTTP access to the multimedia resources. However the diagram illustrates a scenario where the resources may be distributed across one or more other webservers. This architecture means that the resources are not restricted to a single site and may be accessed from other locations. This adds locational transparency and in larger systems allows for load balancing.

6.2 Searching & Rights Checking

Figure 6.2 shows the structure of the webpages in system and how users navigate through them:

- Once a user logs in they are presented with a form where they can enter a keyword to search (Figure 6.2b). After submitting the form the system will query the database (using the JDBC API [20]), checking if the keyword appears in the *title*, *subject* or *description* of any descriptive metadata.

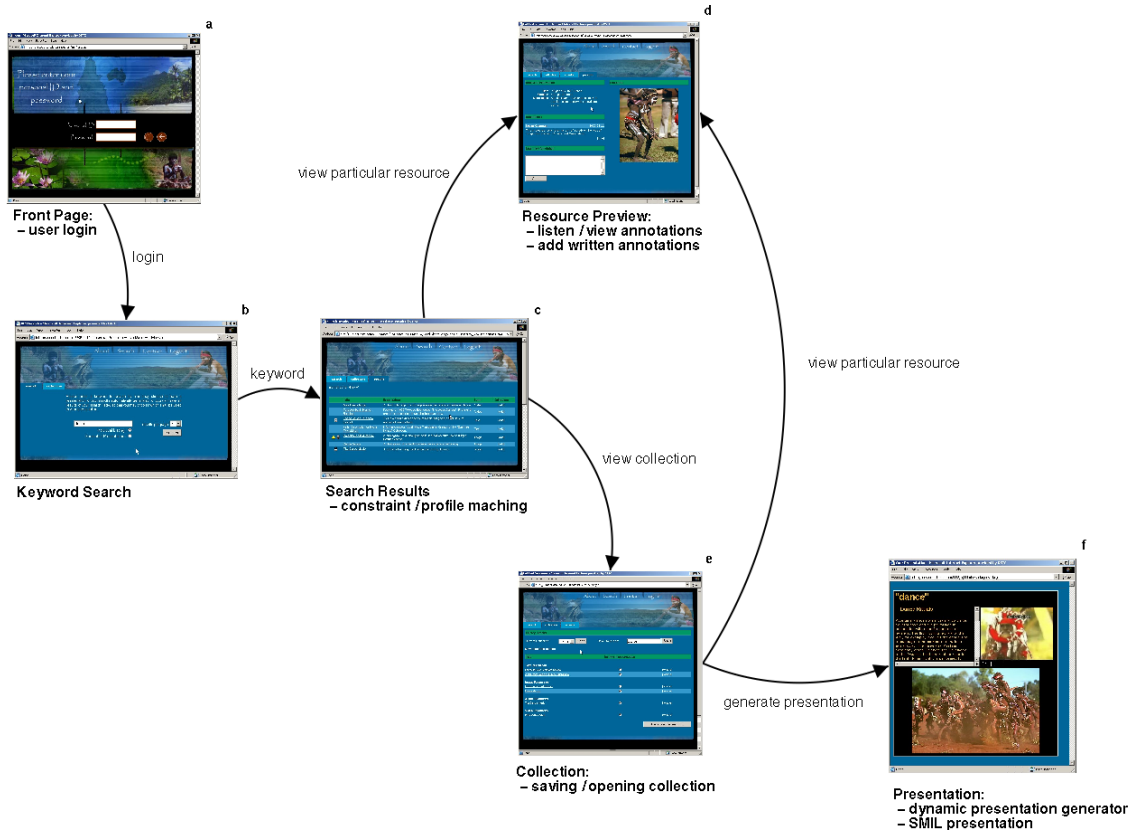

Figure 6.2: User Interaction of the Search, Browse & Retrival Interface

- For each record in the set of results the system will check if there are any rights restrictions associated with the particular resource. If the resource contains restrictions they will be matched against the user's profile and only if the user meets the required criteria will the resource be displayed.

- For each result the *title*, *description* and *type* (text, image, audio or video) is displayed. Figure 6.3 (corresponding to Figure 6.2c) shows the results returned on the keyword "dance". The icons (Figure 6.4) in the far left column have special meaning:

  rights icon signifies that there are rights restrictions on the particular resource. The user has permission to access this resource but should still be made aware that restrictions exist. When the mouse is moved over this icon the rights restrictions are displayed. Even though the user has permission to access this resource it is still important to make them aware that specific restrictions exist. For example when the user is brows-

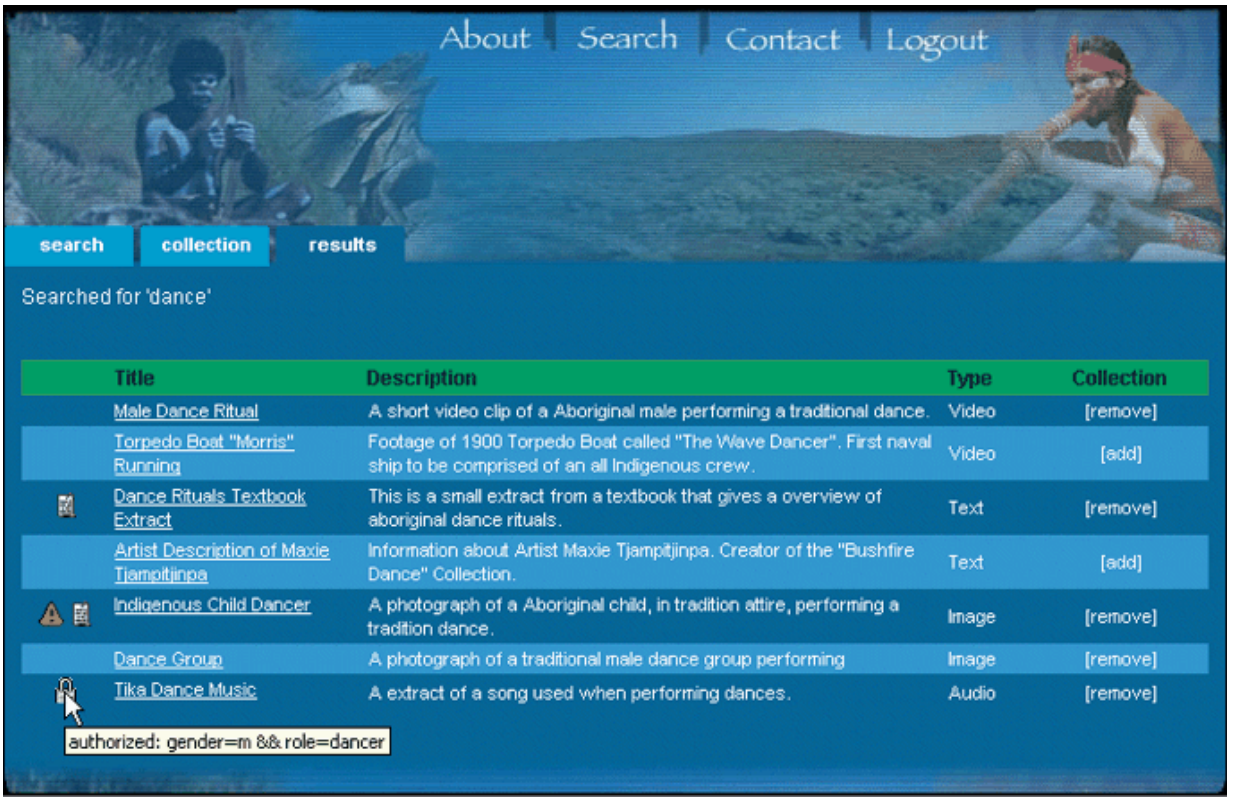

Figure 6.3: Filtered results returned by a search for “dance”

Figure 6.4: *rights*, *warning* and *context* icons

ing the search interface with another person sitting nearby, they should be aware of the nature of the resource before opening it.

warning icon signifies that the particular resource has a warning associated with it. When the mouse is moved over this icon the warning is displayed.

context icon signifies that the particular resource has a contextual constraint associated with it. When the mouse is moved over this icon the text relating to this constraint is displayed.

6.3 Viewing Resources

As seen in Figure 6.3 the *title* of the resource is a hyperlink. Following this link takes the user to the resource preview page (Figure 6.2d). Here the user can view the particular resource and read or listen to any annotations associated with it.

A facility for adding written annotations is also provided so users who do not have access to the Metadata Editor/Generator application can also annotate particular resources.

6.4 Collections

The Search, Browse & Retrieval Interface implements the notion of a "shopping cart", called a *Collection*. As users find resources that are of interest to them they add them to their collection. This allows users to build up collections of items across different searches.

The collections page (Figure 6.2e) allows users to manage their collections. It enables them to:

- view the items that are currently in their collection;
- remove items from their collection;
- save their collection: the list of items in the collection is written to the `COLLECTIONS` table in the database;
- open saved collections stored in the database.

6.5 Multimedia Presentation (SMIL)

Users may select the items in their current collection they would like presented as a dynamically generated SMIL presentation. For each media item a region in which it will be displayed is created. These are created according to three XML templates which define the layout of the presentation. These are shown in Figure 6.5.

The process of creating the presentation is as follows:

- the presentation templates are read in (using the JDOM API [19]) and region sizes adjusted to match those of the media items being displayed;

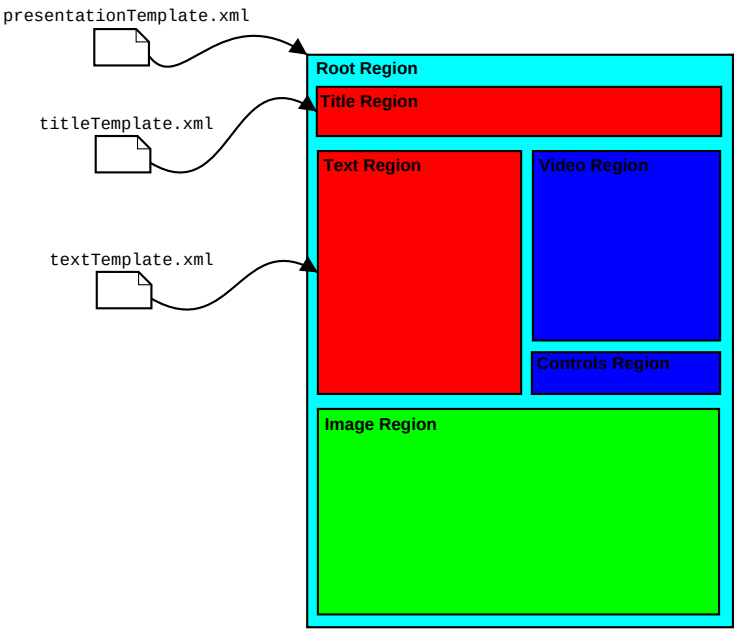

Figure 6.5: SMIL generation presentation layout.

- the media items URL, stored as the Dublin Core *Identifier* in the database is inserted so that it may be fetched when the presentation is run;
- because the most common player for SMIL presentation is *RealPlayer*, and as *RealPlayer* uses the *RealText* format to display text, any textual resources must first be converted from text/HTML to *RealText* format. The `titleTemplate.xml` and `textTemplate.xml` templates are provided to define the *RealText* documents that are created;
- if the presentation contains audio for video media items a *Controls* region is added that contains play, stop and pause buttons;
- the resulting SMIL file is saved to a temporary folder.

The SMIL presentation is then embedded into the presentation page (Figure 6.2e) of the user's browser.

Some users have expressed interest in being able to save SMIL presentations that they find particularly interesting. A hyperlink to the SMIL file is provided at the bottom of the presentation for this purpose.

An example of a presentation of media items for a search for "dance" is shown in Figure 6.6

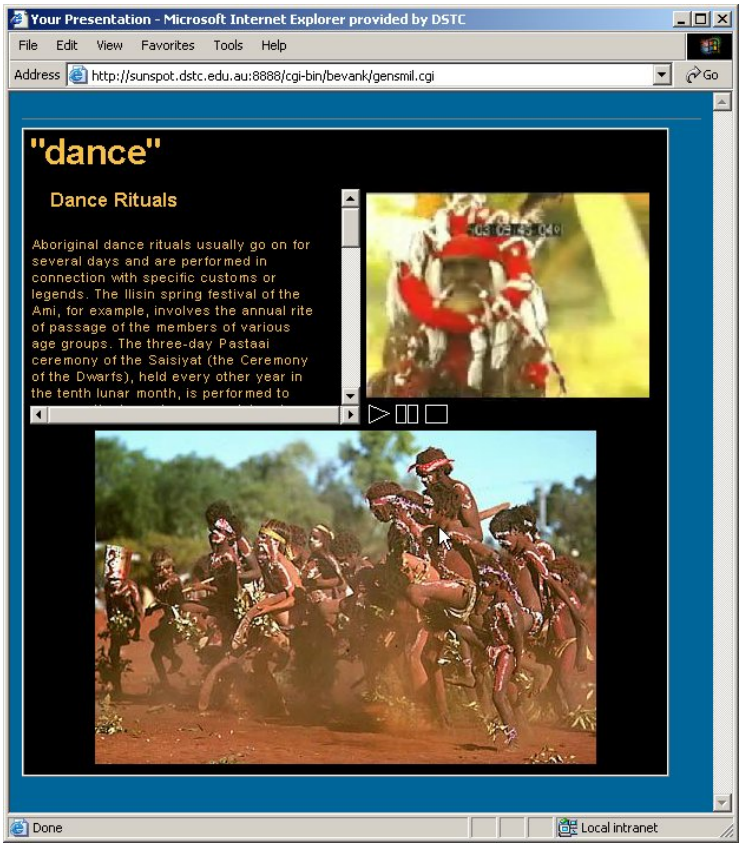

Figure 6.6: SMIL presentation for “dance”.

Chapter 7

Conclusions & Future Work

This chapter discusses the problems associated in dealing with the sensitive nature of Indigenous knowledge and communities. Further research directions are then considered and final thoughts and conclusions presented.

7.1 Problems

The technical aspects of this project were less problematic then the social and political concerns. The use of tried and tested platform independent tools such as Java and JDBC, and flexible data representation mechanisms such as XML and XML Schemas, simplified the task of extending or adapting current technologies to suite the requirements of Indigenous knowledge. The political and social problems that arose during the project proved harder to overcome. Originally the goal was to trial the software with an actual community and collection to gain feedback and evaluate the tools effectiveness. A number of possible collections were identified and attempts were made to collaborate with the relevant owners in order to digitise them. This proved to be a complex and difficult process. The following major problems emerged:

- decisions to grant permission to use a particular collection typically came from groups of elders, committees or community groups rather than from individuals. Gaining consensus from these groups was often a lengthy process;
- collections were often in the care of museums, libraries or other institutions (e.g. Presbyterian Church) and were not in the care of traditional owners. This meant that agreement

had to be reached with both groups. These groups had differing and sometimes conflicting objectives regarding the use of a particular collection;

- the secret and sacred nature of some collections and the fear of exploitation meant that certain groups were apprehensive or refused to allow access to the collection;
- the trust requirements (identified in Section 3.6.1) mean that a trusted person is required to enter and validate users, ensuring user profiles accurately describe the person. Many groups expressed major concerns over having a single person (with some computer literacy skills) having complete control over who is allowed to access the system.

On reflection of these issues and considering the logistics of deploying the software it was decided that the best approach to applying the tools to an actual community would be to make contact with a organisation more adept in dealing with the relevant communities and work with already established Keeping Places. Thus allowing us to focus on the technical aspects and providing us with a single point of contact through which to communicate with the elders or representatives of a community.

7.2 Future Work

The wide variety of different possible usage scenarios for the software means that there is large scope for future research and extensions to the current software.

7.2.1 Annotations

Annotations as Resources

Spoken annotations become a form of resource themselves. In some cases the annotation to a resource is as or more valuable than the resource itself. This also means that some annotations also require the specific rights constraints that are currently applied to normal resources. It may also ensue that users want to add annotations to annotations. In light of this it would be better to treat annotations as a special type of resource that is linked or related to a normal resource.

The ability to attach pre-existing audio or text files to a particular resource as an annotation would also be advantageous. This means that users may attach existing annotations that were not recorded with the Metadata Editor/Generator application.

Regional Annotation

Currently annotations apply to a resource as a whole. Facilities to describe segments or regions of a resource allow for more specifically focussed annotations. Future development on this project will allow users to select a particular region of a image or segment of video or audio file and attach a separate annotation to it. For example if the resource is a photograph containing a number of different people, a user may select each person separately and describe who they are.

7.2.2 Describing 3D Objects

Adding metadata to three dimensional objects is another area for future research. The National Museum of the American Indian has produced a number of QuickTime VR multimedia files of some of the artefacts. Users may rotate and view these objects from different angles and may zoom in on specific regions. The QuickTime VR format simply splices separate two dimensional images together, it is not a proper three dimentional modeling language.

The Web3D Consortium have develed the Extensible 3D (X3D [25]) Graphics specification. X3D defines a Virtual Reality Modeling Language (VRML) using XML. Adding annotations or descriptive metadata to certain regions of these 3D formats is a logical next step in better describing multimedia resources.

7.2.3 Layering on Maps

Indigenous culture and knowledge is closely tied to geography and landscape. By overlaying metadata on maps describing specific areas or geographical features the Indigenous close connection with the land is also represented in the digital domain. Maps may even be used as the search and browse interface. Users will be presented with a map showing clickable “hot spots” where metadata has been added to a specific region or geographical feature.

7.2.4 Security Additions

The security issues that have arisen may be considered in much greater depth. Both the granularity of access control and cryptography of both the resources and metadata are areas for future research and development.

Rights Granularity

The increased granularity provided by allowing regional annotations is also applicable to rights constraints. In some cases different regions of an image or segments of a video resource may require different rights constraints. For example considering the laws of "sorry business" where it is inappropriate to view images of a person who has recently died. If the person is depicted in a particular image then it would desirable to set the rights constraints to the specific region in which the person appears to be different from the rest of the image. The same applies to segments of a video or audio resource than may have different rights constraints.

Resource Encryption

Currently all resources are stored on disk and require the software to perform the necessary rights checks to control access. If the software can be circumvented then the resources may be viewed without enforcing the specified rights constraints. To prevent this the resources themselves need to be protected by encryption.

Simple open source tools exist that may be used with the Java Cryptography Architecture (JCA [26]) to incorporate the encryption of images and text into the existing software. This would done using a public/private key encryption system where only the software tools shall have the keys necessary to decrypt resources. This forces all access to the resources to go through the software, thus ensuring that any rights constraints are evaluated before allowing access to the resource. The encryption of audio and video resources may require more specialised tools as most open source encryption tools available for use with the JCA are aimed at the encryption of images and text.

Watermarking

Watermarking involves adding trademark information to images (or possibly other multimedia resource). It provides a way of ensuring the copyright information always "travels" with the resource itself. Extending the software to allow users to add specific textual or graphical copyright information is another step in ensuring proper attribution to the traditional owner and may help prevent misuse.

7.2.5 XML Databases

In Section 4.3 the reasons for the choice of storing metadata in relational tables was outlined. Consideration of the possible future developments and insight into the vastly different scenarios in which the software may be used, highlights the problem associated with storing the metadata in static relational tables. From a data representation perspective XML is far more suited to the system's requirements. Ideally a storage method is needed that incorporates the flexibility and interoperability of XML and the speed and efficiency of relational database search mechanisms.

Recent work within the XML community has produced a number of native XML databases [27]. These systems allow applications to manipulate and process their data in XML format while still taking advantage of some of the performance and features of more mature relational database systems. Using a native XML database allows the Metadata Editor / Generator application's data model (and therefore its interface) to be completely defined by XML Schemas that are easily customised. In addition the Search, Browse & Retrieval Interface will still provide the efficient searches and other performance improvement techniques provided by relational databases.

7.3 Conclusion

Our ever increasing reliance on digital content has led to a number of new technologies for the management of digital resources. These developments include technologies focusing on enforcing copyright and intellectual property laws to govern usage. These include a number of new rights expression languages such as XrML, ODRL, and the MPEG-21 framework.

At the same time Indigenous communities are becoming aware of the potential benefits which digital technologies can offer them with regard to the documentation and preservation of their histories and culture. These communities are coming to understand the opportunities for misuse and misappropriation of their knowledge which may accompany digitisation, and would like to take advantage of new rights management technologies.

Investigation into the traditional laws and protocols that govern the use of Indigenous knowledge has shown that they are both complex and dynamic. The concept of ownership in Indigenous cultures differs from that of the western notion, with Indigenous cultures having the notion of collective rather than individual ownership. Factors such as a person's gender, status within a tribal unit and relationships to people being represented are all factors determining access to a resource. These factors can differ vastly between communities, making it difficult to define a

standard set of laws and protocols.

New rights management technologies, based on modern intellectual property law fail to meet the complex and dynamic requirements of managing Indigenous knowledge. In this thesis we have extended current rights management technologies and produced a set of open source software tools that provide support for traditional and customary laws and protocols.

The software is provided in two parts with both sharing a common metadata store. The Metadata Editor/Generator Java application allows authorised uses to enter metadata to describe resources. This includes specifying the rights constraints governing usage and adding spoken or written annotations to further describe the resource. Administrator users can add and edit user profiles as well as load and customise the schema defining rights metadata. The Search, Browse & Retrieval Interface was developed as a number of dynamically generated webpages to allow access to the metadata stored by the Metadata Editor/Generator. Authorised users can perform keyword searches and results are filtered according to a users profile and the rights constraints on a resource. Users may build up their own custom collections of interesting resources and have them displayed as dynamically generated multimedia presentations.

The software described in this thesis will hopefully empower the tradition owners of Indigenous knowledge, allowing them to specify ownership and usage and ensuring that traditional laws and protocols are upheld in the digital domain. Through this thesis it is hoped that digital technology, which has brought huge benefits to commercial and research fields, may also be used with the preservation of Indigenous knowledge, cultures and history threatened by globalisation.

Bibliography

[1] XrML. *eXtendible rights Markup Language*. ContentGuard, 2002. http://www.xrml.org.

[2] ODRL. *Open Digital Rights Language*. IPR Systems, 2002. http://www.odrl.net.

[3] MPEG Requirements Group. MPEG-21 Overview V.4. MPEG Document ISO/IEC/JTC1/SC29/WG11/N4801, MPEG Requirements Group, May 2002.

[4] WIPO. Intellectual Property Needs and Expectations of Traditional Knowledge Holder. WIPO Report on Fact-Finding Missions on Intellectual Property and Traditional Knowledge(1998-1999), World Intellectual Property Organisation, Geneva, April 2001.

[5] Hillmann, D. *Using Dublin Core*. Dublin Core Metadata Initiative Recommendation, 12 April 2001. http://www.dublincore.org/documents/2001/04/12/usageguide.

[6] ICOM/CIDOC. CIDOC Conceptual Reference Model. Proposed ISO 21127, ICOM/CIDOC Documentation Standards Group, June 2002. http://cidoc.ics.forth.gr.

[7] Hunter, J. Combining the CIDOC/CRM and MPEG-7 to Describe Multimedia in Museums. In *Proceedings Museums and the Web*, Bonston, April 2002. http://www.archimuse.com/mw2002/papers/hunter/hunter.html.

[8] W3C. *XML Schema Part 0: Primer*. W3C Recommendation, 2 May 2001. http://www.w3.org/TR/xmlschema-0.

[9] W3C. *XML Schema Part 1: Structures*. W3C Recommendation, 2 May 2001. http://www.w3.org/TR/xmlschema-1.

[10] W3C. *XML Schema Part 2: Datatypes*. W3C Recommendation, 2 May 2001. http://www.w3.org/TR/xmlschema-2.

[11] W3C. *XML Encryption Syntax and Processing*. W3C Recommendation, 4 March 2002. http://www.w3.org/TR/xmlenc-core.

[12] W3C. *XML Signature Syntax and Processing*. W3C Recommendation, 7 August 2002. http://www.w3.org/TR/xmlsig-core.

[13] W3C. *Synchronized Multimedia Intergration Language (SMIL2.0)*. W3C Recommendation, 7 August 2001. http://www.w3.org/TR/smil20.

[14] Githaiga, J.W. Intellectual Property Law and the Protectuib of Indigenous Folklore and Knowledge. *Murdoch University Electronic Journal of Law*, 5(2), 2 June 1998.

[15] Janke, T. Our Culture Our Future: Indigenous Cultural and Intellectual Property Rights. Report on Australian Indigenous Cultural and Intellectual Property Rights, Michael Frankel & Company, 25 November 1999. http://www.icip.lawnet.com.au.

[16] Hunter, J. Rights Extensions for the Protection of Indigenous Knowlege. In *Proceedings Eleventh International World Wide Web Conference*, Honolulu, Hawaii, May 2002. http://www2002.org/CDROM/alternate/748/index.html.

[17] Hunter, J and James, D. The Application of an Event-Aware Metadata Model to an Online Oral History Project. In *Proceedings Fourth European Conference on Research and Advanced Technology for Digital Libraries*, Lisbon, Portugal, September 2000. http://link.springer.de/link/service/series/0558/papers/1923/19230291.pdf.

[18] JMF. *Java Media Framework version 2.1.1a*. Sun Microsystems, 2002. http://java.sun.com/products/java-media/jmf.

[19] JDOM. *Java Document Object Model Application Programming Interface*, 2002. http://www.jdom.org.

[20] JDBC. *Java Database Contectivity (JDBC)*. Sun Microsystems, 2002. http://java.sun.com/products/jdbc.

[21] MySQL. *MySQL Reference Manual for version 4.0.4*. MySQL AB, 2002. http://www.mysql.org.

[22] JSP. *JavaServer Pages (JSP)*. Sun Microsystems, 2002. http://java.sun.com/products/jsp.

[23] K. Dunman. New methods in the care of american indian artifacts, Fall 2001. http://www.focusanthro.org/american_indian_artifacts.htm.

[24] Freier, A.O., Karlton, P., and Kocher, P.C. *The SSL Protocol Version 3.0*. Transport Layer Security Working Group, 18 November 1996. http://wp.netscape.com/eng/ssl3/draft302.txt.

[25] X3D Graphics Working Group. Extensible 3D (X3D) Graphics Working Groups. International Standard ISO/ICE CD 19775:200X, Extensible 3D Graphics Working Group, October 2002. http://www.web3d.org/TaskGroups/x3d/X3DSpec_CD_Preview/index.html.

[26] JCA. *Java Cryptography Architecture (JCA) API Specification & Reference*. Sun Microsystems, December 1999. http://java.sun.com/products/jdk/1.2/docs/guide/security/CryptoSpec.html.

[27] XML:DB. XML:DB Initiative for XML Databases, 2002. http://www.xmldb.org.

Appendix A

XML Configuration File

Below is an example of the XML configuration used by the Metadata Editor / Genterator. A description of this file can be found in Section 5.1.

```
<?xml version="1.0" encoding="UTF-8"?>
<irm>
  <dbstring>jdbc:mysql://server.dstc.edu.au/irm-dbase</dbstring>
  <username>meg-application</username>
  <password>meg-passwd</password>
  <resources>/www/htdocs/irm/resources</resources>
  <access>http://server.dstc.edu.au/irm/resources</access>
  <schema>/home/user/my_schema.xml</schema>
</irm>
```

Appendix B

Rights Schemas

Below is an example of the XML schema file used by SchemaManager component of the Meta-data Editor / Genterator. The SchemaManager is destribed in Section 5.3.

```
<xml version="1.0" encoding="UTF-8"?>
  <schema>
    <rights>
      <gender enabled="true" type="disjoint">
        <vocab>m</vocab>
        <vocab>f</vocab>
      </gender>
      <tribe enabled="true" type="value" />
      <clan enabled="true" type="value" />
      <band enabled="true" type="value" />
      <family enabled="true" type="value" />
      <role enabled="true" type="intesect">
        <vocab>artist</vocab>
        <vocab>musician</vocab>
        <vocab>dancer</vocab>
        <vocab>midwife</vocab>
        <vocab>hunter</vocab>
      </role>
      <status enabled="true" type="intesect">
        <vocab>elder</vocab>
```

```
        <vocab>initiated</vocab>
        <vocab>unitiated</vocab>
        <vocab>child</vocab>
      </status>
    </rights>
  </schema>
```

Appendix C

SMIL

Example of a SMIL file dynamically generated by the Search, Browse & Retrieval Interface. The procoss of generating this file is described in Section 6.5.

```
<smil>
 <head>
  <meta name="copyright" content="IKMTools" />
  <layout>
   <root-layout width="545" height="543" background-color="black" />
   <region id="title" left="5" top="0" width="600" height="50" />
   <region id="text" left="5" top="50" width="280" height="200" fit="scro
   <region id="video" left="290" top="50" width="250" height="180" fit="m
   <region id="images" left="60" top="255" width="425" height="283" fill=
   <region id="buttons" left="290" top="230" width="250" height="20" fit=
  </layout>
 </head>
 <body>
  <par>
   <text src="title.rt" region="title" begin="0s" dur="20s" />
   <text src="text.rt" region="text" begin="0s" dur="20s" />
   <video src="http://irm.dstc.edu.au/resources/000010.mpeg"
          region="video" clip-begin="0s" dur="20s" />
   <img src="http://irm.dstc.edu.au/resources/buttons.gif"
        region="buttons" begin="0s" dur="20s">
```

```
    <anchor href="command:play()" target="_player" coords="3,2,20,20" />
    <anchor href="command:pause()" target="_player" coords="23,2,40,20" /
    <anchor href="command:stop()" target="_player" coords="50,2,67,20" />
   </img>
   <audio src="http://irm.dstc.edu.au/resources/000012.wav"
          clip-begin="0s" dur="20s" />
   <img src="http://irm.dstc.edu.au/resources/000004.jpg"
        region="images" begin="0s" dur="20s" />
  </par>
 </body>
</smil>
```

Appendix D

SQL Create Statments

```
CREATE TABLE DC (
       id mediumint not null auto_increment,
       resource_id VARCHAR(20) not null,
       title VARCHAR(200),
       creator VARCHAR(200),
       subject VARCHAR(200),
       description LONGTEXT,
       publisher VARCHAR(200),
       contributor VARCHAR(200),
       date_created DATE,
       date_issued DATE,
       type VARCHAR(10),
       format VARCHAR(5),
       identifier VARCHAR(200),
       source VARCHAR(200),
       language VARCHAR(20),
       relation VARCHAR(200),
       coverage VARCHAR(200),
       primary key(id)
);

CREATE TABLE PERSON (
```

```
        login_id VARCHAR(20) not null,
        name VARCHAR(200) not null,
        type VARCHAR(20),
        password VARCHAR(20) not null,
        organisation mediumint references ORGANISATION,
        tribe VARCHAR(200),
        clan VARCHAR(200),
        band VARCHAR(200),
        family VARCHAR(200),
        birth_place VARCHAR(200),
        birth_date DATE,
        death_date DATE,
        gender ENUM('m', 'f', 'u'),
        status VARCHAR(200),
        role VARCHAR(200),
        mother VARCHAR(200),
        father VARCHAR(200),
        spouse VARCHAR(200),
        primary key(login_id)
);

CREATE TABLE PERSONGROUP (
        group_id mediumint not null,
        name VARCHAR(200),
        member VARCHAR(20) not null references PERSON,
        primary key(group_id)
);

CREATE TABLE ORGANISATION (
       organisation_id mediumint not null,
       name VARCHAR(200),
       primary key(organisation_id)
);
```

```
CREATE TABLE EVENT (
       event_id mediumint not null auto_increment,
       event_type ENUM('dc', 'rights', 'annotations', 'tc',
                       'login', 'user_account', ) not null,
       date TIMESTAMP not null,
       place VARCHAR(200),
       input_id VARCHAR(200) not null references DC,
       output_id mediumint not null references RIGHTS,
       agent_id VARCHAR(20) not null references PERSON,
       role ENUM('owner', 'admin', 'viewer'),
       primary key(event_id)
);

CREATE TABLE RIGHTS (
       rights_id mediumint not null auto_increment,
       resource VARCHAR(20) not null references DC,
       user VARCHAR(20) not null references PERSON,
       constraint_type ENUM('view', 'copy', 'handle',
                            'warning', 'request') not null,
       constraint_value LONGTEXT not null,
       primary key(rights_id)
);

CREATE TABLE ANNOTATIONS (
       annot_id mediumint not null auto_increment,
       resource VARCHAR(20) not null references DC,
       comment LONGTEXT not null,
       primary key(annot_id)
);

CREATE TABLE TRIBAL_CARE (
        resource VARCHAR(20) not null references DC,
```

```
        type enum('orientation', 'storage', 'preservation') not null,
        value LONGTEXT,
        primary key(resource, type)
);

CREATE TABLE OWNERSHIP (
       resource VARCHAR(20) not null references DC,
       owner VARCHAR(20) not null references PERSON,
       primary key(resource, owner)
);

CREATE TABLE COLLECTION (
        name varchar(200) not null,
        user varchar(20) not null references PERSON,
        resource mediumint not null references DC,
        primary key(name, user, resource)
);
```

Appendix E

Co-authored Paper: *Software Tools for Indigenous Knowledge Management*

The following paper contains silimar content to this thesis. It presents some background to the work done on this project and describes the set of software tools that have been developed.

The authors of the paper are Jane Hunter, the supervisor of this thesis; Bevan Koopman, the author of this thesis; and Jane Sledge, the Information Resource Manager of the National Museum of the American Indian, who also provided input and ideas though a collaborative project.

The paper is aimed at both a technical and non-technical audience as it is applicable to readers from the museum, cultural institution and Indigenous background, as well as readers from metadata, multimedia and digital library disciplines.

Software Tools for Indigenous Knowledge Management

Jane Hunter[1], Bevan Koopman[2], and Jane Sledge[3]

[1] DSTC, Brisbane jane@dstc.edu.au
[2] ITEE Dept. University of Queensland bevank@dstc.edu.au
[3] Smithsonian NMAI, Washington sledgej@nmaicrc.si.crc

Abstract. Indigenous communities are beginning to realize the potential benefits which digital technologies can offer with regard to the documentation and preservation of their histories and cultures. However they are also coming to understand the opportunities for misuse and misappropriation of their knowledge which may accompany digitization. In this paper we describe a set of open source software tools which have been designed to enable indigenous communities to protect unique cultural knowledge and materials which have been preserved through digitization. The software tools described here enable authorized members of communities to: define and control the rights, accessibility and reuse of their digital resources; uphold traditional laws pertaining to secret/sacred knowledge or objects; prevent the misuse of indigenous heritage in culturally inappropriate or insensitive ways; ensure proper attribution to the traditional owners; and enable indigenous communities to describe their resources in their own words. Hopefully the deployment of such tools will contribute to the self-determination and empowerment of indigenous communities through the revitalization of their cultures and knowledge which have been eroded by colonization, western laws, western cultures and globalization.

1 Introduction

Using multimedia technologies, indigenous groups have been able to record and preserve significant aspects of their cultures including languages [1], ceremonies, dances, songs, stories, symbols, design, artwork, tools, costumes, historical photographs, film, videos and audio tapes [2,3,4]. Documentation of indigenous knowledge and history has become an extremely important tool to ensure the survival and self-sustainability of indigenous tribes and cultures, and to provide evidence of past injustices and to support claims of original ownership.

Although digitization is ideal for sharing, exchanging, educating and preserving indigenous cultures, it also creates ample opportunities for illicit access to and misuse of traditional knowledge. It is essential that traditional owners are able to define and control the rights and access to their resources, in order to: uphold traditional laws; prevent the misuse of indigenous heritage in culturally inappropriate or insensitive ways; and receive proper compensation for their cultural and intellectual property. Finally it is essential that indigenous communities are able to describe and contextualize their culturally and historically significant collections, in their own words and from their own perspectives.

In this paper we have investigated the application of IT security mechanisms to the rights management of indigenous collections. In consultation with staff from the NMAI Cultural Resources Center [5], we have developed a set of low cost, simple-to-use and robust software tools designed to enable the description, annotation and rights management of collections of mixed-media digital and physical objects belonging to indigenous communities. We have also developed a search, retrieval and presentation interface which retrieves different result sets, depending on the user's profile, and aggregates the results automatically into coherent multimedia (SMIL [6]) presentations.

Because of the enormous diversity of indigenous cultures, the system has been designed so that it can easily be customized to support the unique requirements of specific communities. The immediate future involves working closely with a small number of indigenous communities to refine the software to suit their specific requirements and projects. Eventually we hope to make the software downloadable from the Internet and freely available to indigenous communities for non-profit use and to provide training in its use and maintenance.

2 Background and Requirements

The work described here began with an investigation of the ability of current and evolving information technology tools for rights management (e.g., frameworks, markup languages, metadata models and standards) to support the unique needs of indigenous communities. We found that initiatives such as MPEG-21 (Multimedia Delivery Framework) [7] and XrML [8] are primarily concerned with e-commerce and protecting the commercial rights of content owners. They are built on the premise of modern intellectual property law regimes and the notions of individual property ownership for a limited duration, which are alien and detrimental to indigenous cultures. MPEG-21 and XrML do not support the specific requirements expressed by indigenous communities which are needed to protect indigenous knowledge or enforce tribal customary laws.

In an earlier paper [9], specific extensions to XrML, in the form of customary constraints, were developed to support the description of customary or traditional laws which commonly affect access. An analysis of tribal laws across Aboriginal and Native American communities revealed that the following common factors or variables which may determine access to traditional knowledge:

- native/non-native restrictions;
- the user's membership of a particular clan or tribe;
- the user's status within the tribe (e.g., elder, initiate, child);
- the user's role within the tribe (e.g., dancer, artist, midwife, healer);
- the user's gender (male, female);
  - moon (menstrual cycle) restrictions;
  - pregnancy restrictions;
- the relationship of the viewer/user to the people, animals or objects depicted in the resource;
- the death of people recorded in a resource;
- human remains access restrictions - no access should be provided to images of human remains, or specifics about the disposition of human remains repatriated to tribal communities;
- the context in which the resource will be reused or reproduced.

Staff at the NMAI CRC also expressed the need to be able to document traditional care information associated with culturally significant physical artefacts in their custody. Kristina Dunman has described the meaning and importance of traditional care of American Indian artefacts in [10]. Jim Pepper Henry (Repatriation Program Manager at the NMAI) also provided the following list of traditional care constraints which are frequently requested by American Indian tribal representatives on the storage and handling of objects:

– directional orientation;
– segregation from other objects or other tribes' materials;
– storage on higher shelves, use of wooden shelving;
– cover from view with cloth or muslin;
– ceremonial feeding/dusting with corn pollen and/or other materials;
– cleansing with smoke (cedar, sage, tobacco, sweetgrass etc.);
– freezing restriction;
– low oxygen restriction (no CO2 bubble or covering with plastic or other non permeable materials);
– sacred/ceremonial bundle dis-assemblage restriction (bundle to remain intact and unopened);
– ceremonial pipe assemblage restriction (bowl and stem should be housed disconnected).

Similar traditional care recommendations also apply to culturally sensitive artefacts belonging to Aboriginal and Torres Strait Islander communities, which are held in museums, archives, cultural centres and *keeping places*.

The software should also enable the traditional owners to describe, contextualize and annotate resources in their own words, their own languages and from their own perspectives. The importance and value associated with enabling spoken annotations (in addition to textual annotations) was multi-faceted:

– Spoken annotation tools reinforce and support the oral tradition which is so strong in many indigenous cultures;
– Spoken annotations are an easier and more natural interface for user input than keyboards, particularly for communities with low computer literacy and poor keyboard skills;
– Spoken annotations represent new language resources which can be used to help preserve threatened languages;
– Photos and videos can act as a trigger for the Indigenous elders to record their stories, as spoken annotations to the visual resources.

In addition, users should be able to to view or listen to associated annotations which are clearly attributed to individuals. This approach supports the unambiguous documentation of all views/perspectives - even if they are different or contradictory. The software explicitly displays "who said what and when" rather than displaying only the view of a Museum Curator which may have been deduced from a number of different sources of varying reliability.

Figure 1 below illustrates how we envisage the software could be used to assist with the protection, preservation and repatriation of indigenous knowledge and artefacts which are being shared, exchanged or returned from museums, archives, private anthropologists' collections and cultural institutions back to their original owners. (N.B. We are not considering the repatriation of human remains within the scope of this project.)

No single approach is applicable to the repatriation of indigenous information, knowledge and/or artefacts. For example, Australian law differs from US law and additionally, each tribal community will have its own unique needs and requests. In the United States, while the Native American Graves Protection and Repatriation Act (NAGPRA) specifies the types of objects and sites to be protected and/or repatriated, it is expected that some tribal communities will want

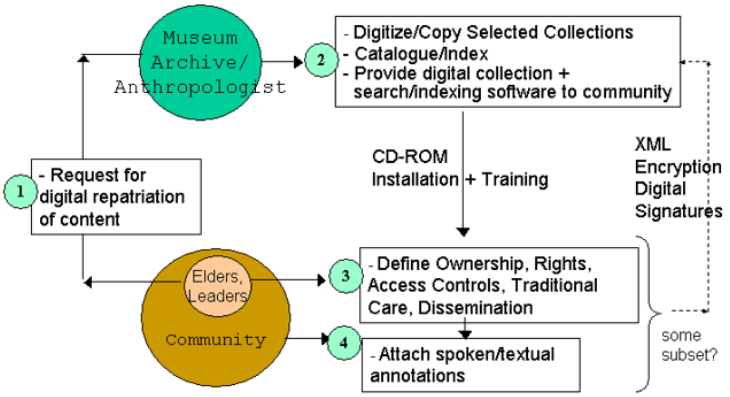

Fig. 1. Software Usage Scenario

access to the records of all objects in museum collections associated with their community and be satisfied with digital surrogates and access to physical objects when requested. Going beyond the requirements of NAGPRA, the NMAI has established a Culturally Sensitive Collections Care Program to respond to areas of concern of Native peoples with regard to the maintenance, presentation, and disposition of sensitive materials and information in the collections of the museum. Jim Pepper Henry, Assistant Director for Community Services at NMAI says: ”This Program is to be implemented with regard to the wishes and concerns of indigenous communities and traditional leaders and structured within the boundaries of the obvious and reasonable limitations of the institution. The basis for this program is formed through consultations with official tribal representatives, tribal elders and traditional leaders, and museum staff with respect to museum policies and procedures, and US federal legislation including the National Museum of the American Indian Act, American Indian Religious Freedom Act (AIRFA), Native American Graves Protection and Repatriation Act (NAGPRA), and the Collections Policy of the National Museum of the American Indian.The Program Committee is cognizant of the fact that it is operating within the constraints of a national museum, and acknowledges that in some instances, dependant upon the beliefs and feelings of the Native community concerned, the proposed implementation of sensitive collections care practices within the museum setting may be deemed inappropriate or presumptuous. In such circumstances, the Committee will endeavor to abide by any alternative direction offered by the concerned Native group.”

The aim of the software described here is to support such a Program.

The envisaged usage and application of the software as described above, introduced certain requirements and design constraints which needed to be satisfied. These included:

– Security mechanisms - because of the sacred/secret nature of the content with which we are dealing, it is essential that the IT security mechanisms which are employed are impenetrable and reliable;
– Simple user interfaces - many of the potential users of this system will have low computer literacy, so simple intutive user-friendly interfaces are essential;

– Robustness - the system must be able to stand up to the rigours of unexpected input by users with little prior computing experience;
– Low cost - in order to make the software open source and accessible to indigenous and grass-roots communities, it must be built as inexpensively as possible, using tools which are ideally free;
– Interoperability - the software tools should be built on international standards - Dublin Core [11], CIDOC CRM [12], MPEG-21 [7], XrML [8]- in order to ensure maximum interoperability between disparate databases;
– Portability - it should be able to run on a range of platforms and operating systems. Java (JDBC, JSP), XML and SMIL have been used as the software development environment to ensure transparent portability across platforms;
– Flexibility - The customary laws and intellectual property needs of traditional knowledge holders vary enormously among indigenous communities throughout the world. Quite often the views within a single clan can vary significantly and they may also vary over time. Our system attempts to support the common notions associated with traditonal laws within Indigenous communities. In addition, we have provided Schema editing tools in order to provide maximum flexibility and to enable easy customization of the software.

Although a number of other projects have investigated the application of information technology tools to the protection and management of indigenous collections according to customary laws [19-22], they have not approached the problem using international metadata standards or developed generic yet flexible systems which are capable of supporting indigenous communities globally, but are easily customized to support the particular local customs.

The remainder of this paper describes the software which has been developed to support the requirements specified above.

3 System Architecture and Components

This section describes the architecture of the software system, its seperate components and how they fit into the overall process/workflow described in Figure 1.

The system consists of three major components:

1. The Metadata Editor/Generator;
2. The Database;
3. The Search, Retrieval and Presentation Interface.

Figure 2 illustrates the interfaces to these components and the technologies used to build them and integrate them into a single coherent system.

3.1 The Metadata Editor/Generator

This component of the system enables users to input the descriptive, rights and tribal care metadata associated with the objects (either physical or digital) and to attach spoken or written annotations to specific objects.

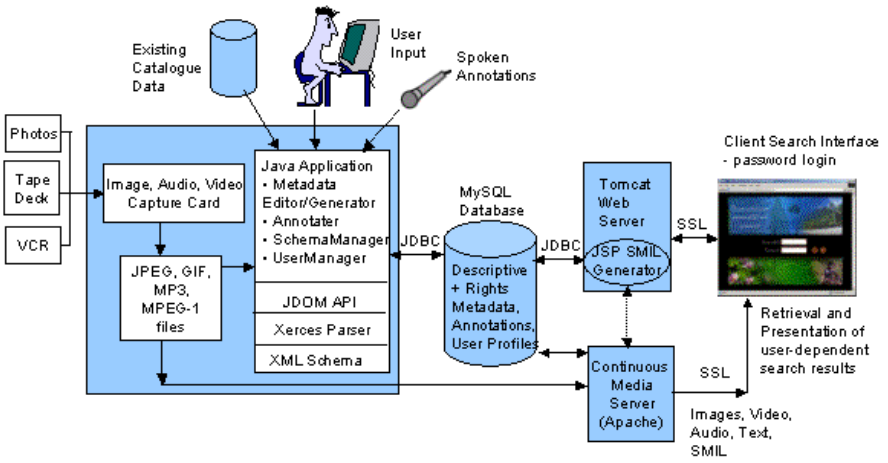

Fig. 2. System Architecture and Workflow

Users require a login ID and password to run this software component and depending on their privileges/user profile, may only be permitted access to certain functionality.

Figure 3 illustrates the three types of metadata which can be input and the users/groups who we envisage will provide them.

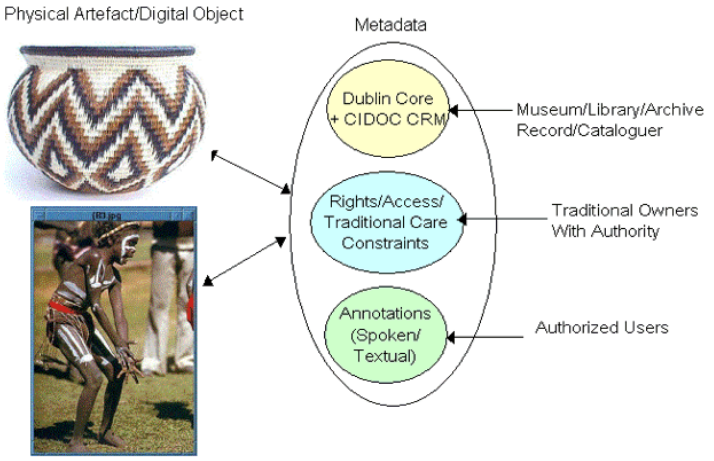

Fig. 3. Metadata Types and Sources

The Descriptive Metadata consists of Dublin Core [11] plus some additional optional elements from the CIDOC CRM [12] for describing physical museum objects, such as material, technique and dimensions. Figure 4 shows the user interface for inputing or editing descriptive

metadata. Alternatively, some or all of the descriptive metadata could be acquired from existing museum/library/archive database cataloguing information.

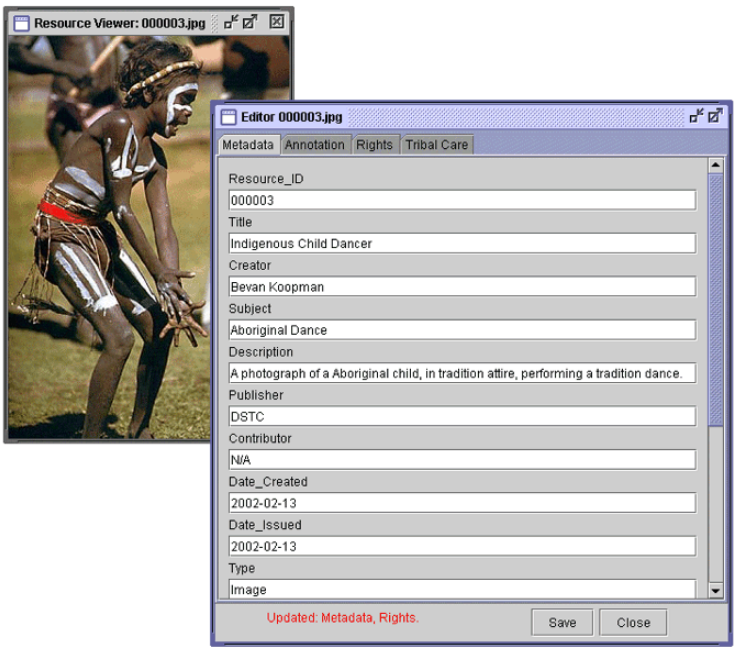

Fig. 4. Descriptive Metadata Input

The Rights Metadata will be provided by the traditional owners or elders of the indigenous community to whom the resource belongs. Only those users with the required access privileges will be able to input or edit the rights metadata. Support has been provided to enable the definition and application of restrictions based on:

- the user's membership of a particular clan or tribe;
- the user's status within the tribe;
- the user's role within the tribe;
- the user's gender;
- the relationship of the user to people, animals or objects depicted in the resource;
- the death of people recorded in a resource;
- the context in which the resource will be reused or reproduced.

Figure 5 shows the user interface for defining access restrictions which depend on the users' tribal affiliation, gender, role and status.

In addition to the descriptive and rights metadata, annotation tools which enable indigenous communities to describe resources in their own words have been provided. Authorized users can input, record and attach either textual or spoken annotations to specific objects or resources. The ability to enter spoken annotations is especially useful and valuable - it provides a natural user interface which supports the oral tradition and allows users to express their stories in their

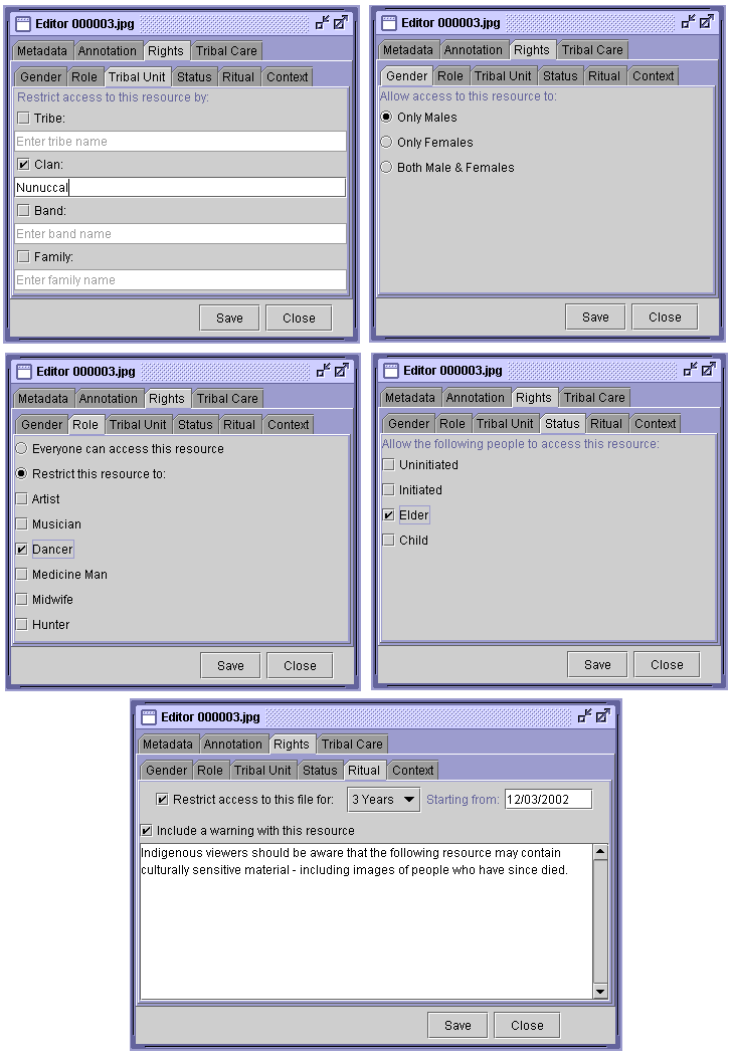

Fig. 5. User Interface for Restricting Access According to Customary Laws

own words and languages. It also acts as a trigger for generating new knowledge and enhancing existing knowledge. Figure 6 below, illustrates the annotation interface. Users can also browse the list of clearly attributed annotations and view/listen to who said what and when about a particular resource.

In response to a request from both museum staff and respresentatives of Indigenous communities, we have also added support for the specification of Tribal Care recommendations. The Tribal Care metadata tool enables Indigenous communities to define the ways in which culturally significant or sensitive physical artifacts should be treated - ways which are acceptable to the cultures represented. Elders are able to define particular spatial orientation, storage and preservation recommendations to ensure the respectful tribal care of physical artefacts. While the actual implementation of traditional care specifications may be difficult, many museums are attempting to

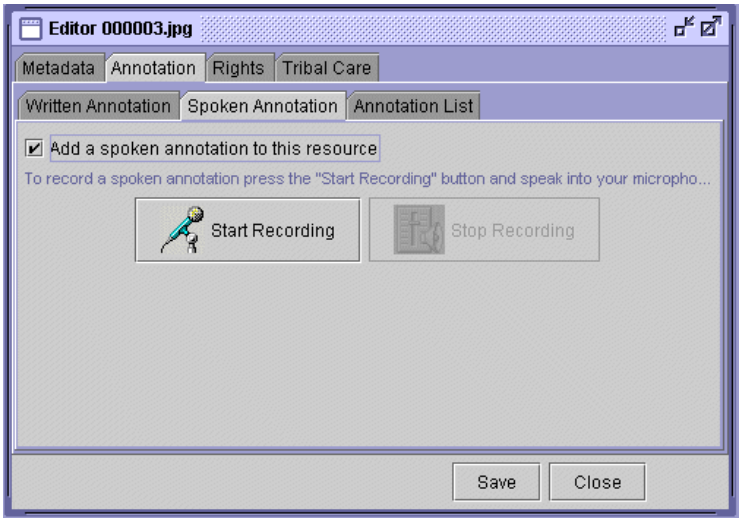

Fig. 6. Annotations

integrate Indigenous beliefs and requests into museum practices and our software will hopefully facilitate this process.

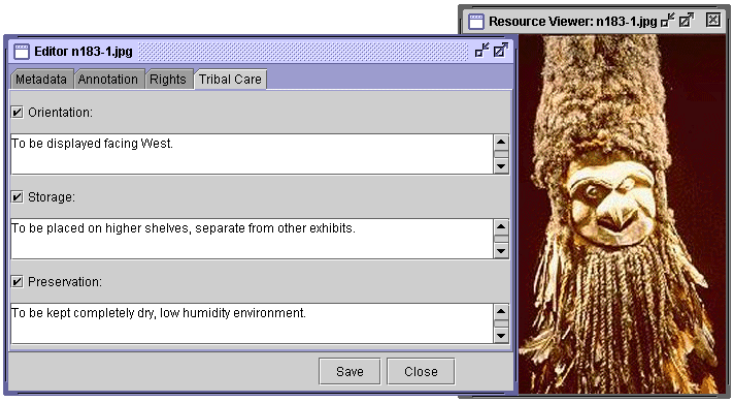

Fig. 7. Tribal Care Specifications

Where possible, the specified access restrictions will be enforced by matching them against the profile of the user. Every user of the system requires a login ID and password. When being allocated a user ID, users will also need to provide supporting documentation to prove their claims of tribal affiliation, status etc.

The User Manager software component, illustrated in Figure 8 allows the systems administrator to add or remove users and to edit their profiles. Authorized elders will be able to enter the rights constraints and authorized users will be able to attach annotations. Hence the software

not only effects what resources users can access and view but also what metadata tools they can access and hence what metadata they can enter.

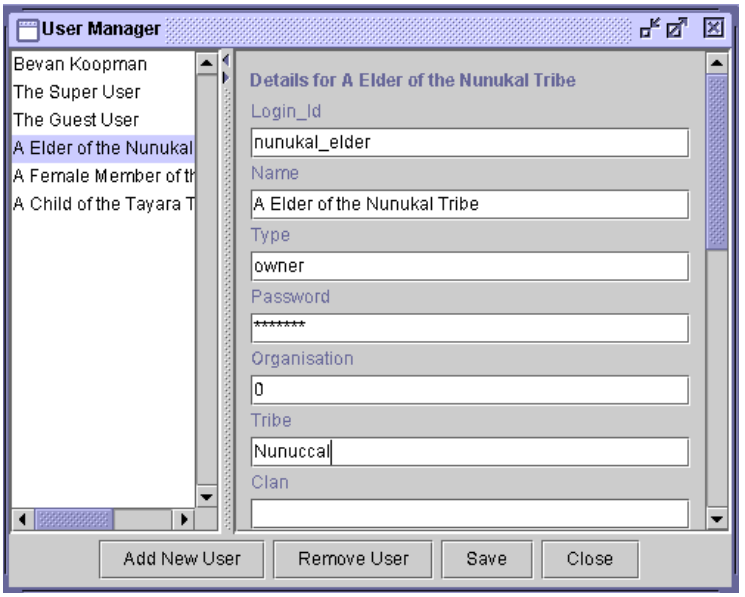

Fig. 8. User Manager

Because of the enormous diversity of indigenous cultures, the system has been designed so that it can easily be customized to support the unique requirements of different communities. Customization is carried out through the SchemaManager tool. Community elders can add new constraints, or remove or refine existing constraints, depending on the traditional laws of their community. An XML Schema [13] is saved to reflect their particular metadata requirements and rights constraints and the user interface is then generated from the saved/selected schema. Figure 9 illustrates the Schema Manager user interface.

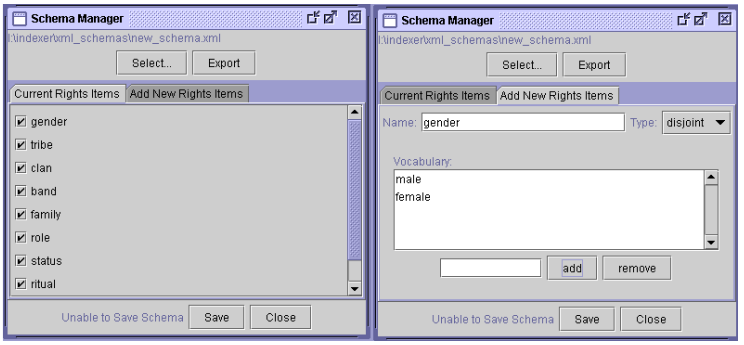

Fig. 9. Schema Manager

3.2 The Database

Saved metadata is stored in relational tables in a MySQL database which is connected through a JDBC (Java Database Connectivity) API [14]. It is envisaged that, because of the sensitivity of the data, this database will not be accessible through the Internet. It will be stored on CD-ROM or the hard disk of a stand-alone supervised workstation, either within a cultural centre, keeping place or a supervised building (such as a library or a school) within the community.

In addition to the metadata which is explicitly saved to the database through the user interface, meta-metadata is also recorded - all changes to the metadata, who was responsible and the date/time of the changes are recorded within the database. This represents an important component of the system's built-in security framework.

3.3 The Search, Retrieval and Presentation Interface

A search, browse and retrieval interface to the collection was built using standard Web Browser technologies (Internet Explorer, Netscape) for the user interface. The advantages of using standard Browsers for the search interface, are their familiarity and widespread availability and the lack of re-engineering necessary should collections eventually be disseminated over wider networks. To access the collection, users must have been allocated a login ID and password and a user profile. The steps below illustrate the typical procedure which users would follow when searching and browsing an indigenous collection:

1. A user logs onto the system using a secure password. Associated with each user is an authenticated user profile which includes information such as tribal/western names, native/non-native heritage, tribal/clan membership, gender, status, role, etc.;
2. The user performs a search on a particular topic e.g., dance (See Figure 10);
3. The software then searches the title, subject, and description metadata associated with each object in the collection, for the specified search term (e.g., dance);
4. For those objects' whose metadata matches the search term, the software compares the objects' rights constraints with the user's profile to determine whether or not the user is permitted to access this object. If they are, then this object will be added to the result set;
5. The list of results/objects which match the search term and which the user is permitted to access, is then displayed - along with any rights constraints, which appear as icons (see Figure 11);
6. Users can click on individual objects to view/play the object and to view the metadata details and any annotations;
7. Users can select objects of particular interest to them and add them to their own collections;
8. The software automatically aggregates those mixed-media objects selected by the user (images, audio clips, video clips, text), and dynamically generates a SMIL (Synchronized Multimedia Integration Language) [6] presentation which is delivered to the user.(see Figure 12);

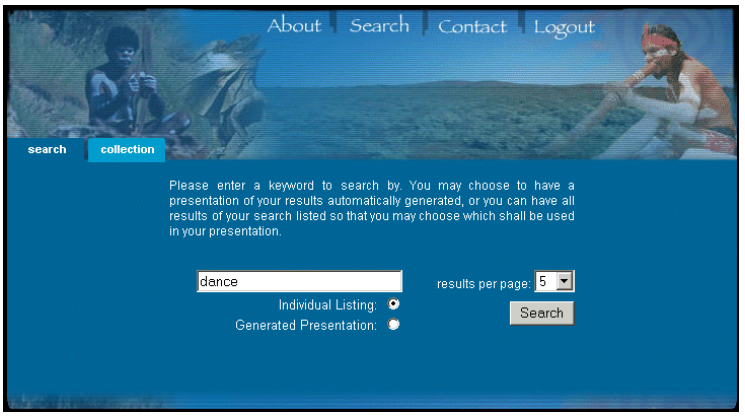

Fig. 10. Search Interface

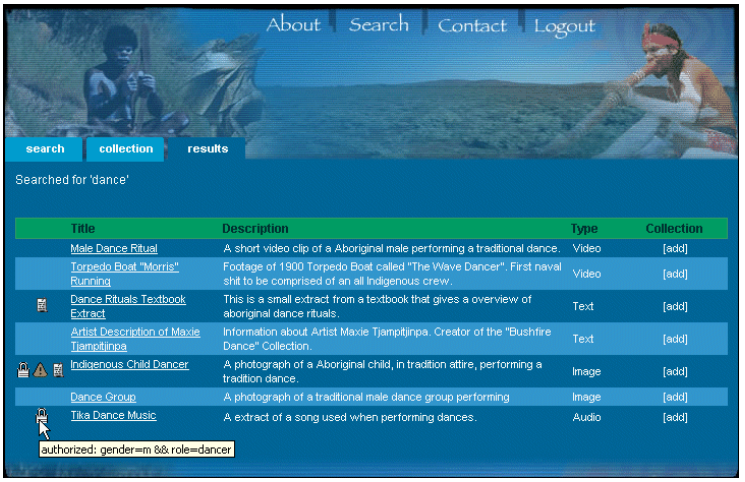

Fig. 11. Search Results

Fig. 12. Multimedia Presentation of Results

4 Future Work and Conclusions

In this paper we have described a software system which has been developed as a result of consultation with representatives from Indigenous communities and staff from museums, archives, libraries and cultural centres in both Australia and North America. However at this stage the software remains relatively untested within real world applications or real communities. Hence the immediate future involves working closely with a small number of indigenous communities to determine:

- whether the software can satisfy the unique requirements of particular Indigenous communities for the management of their culturally sensitive collections or whether further extensions/refinements/modifications will be required;
- whether the dynamic political, social and trust issues (e.g., lack of agreement on access rules, validation of individuals' claims of authority, authorizations within a tribe or community) are greater than the technical problems associated with enforcing them;
- the proper procedures required to enable the succesful and beneficial application of the software to the preservation, description, protection and annotation of indigenous cultural collections by the traditional owners;
- the types of scenarios, situations, collections and communities to which the software is most suited and of maximum benefit;
- whether additional security mechanisms such as XML Encryption [15], XML Digital Signatures [16], SAML (Security Assertion Markup Language) [17], SSL (Secure Sockets Layer) [18] and watermarking techniques, would be applicable and could be trusted to ensure secure access to and transfer of sacred/secret data over networks between distributed remote locations and a common centralized repository.

Given a positive response from the community trials, we plan to make the software freely available and downloadable from the Internet for non-profit use by indigenous communities and to provide training in the usage, refinement and maintainence of the system to interested groups.

Acknowledgements

The work described in this paper has been carried out as part of a Queensland Smithsonian fellowship, funded by the Queensland Government and a collaboration between the Distributed Systems Technology Centre and the Smithsonian's National Museum of the American Indian Cultural Resources Centre (NMAI CRC).

References

[1] First Voices Project, `http://www.firstvoices.com`

[2] NMAI Exhibits online, `http://www.conexus.si.edu`

[3] Barani - Indigenous History of Sydney City,
`http://www.cityofsydney.nsw.gov.au/barani/main.html`

[4] Diwurruwurru - ”Our Message Stick to the World”,
http://arts.deakin.edu.au/Diwurruwurru

[5] National Museum of the American Indian(NMAI), http://www.nmai.si.edu/

[6] Synchronized Multimedia Intergration Language (SMIL), W3C Recommendation, 07 August 2001, http://www.w3.org/TR/smil20/

[7] MPEG-21 Overview,
http://mpeg.telecomitalialab.com/standards/mpeg-21/mpeg-21.htm

[8] XrML, eXtensible rights Markup Language, http://www.xrml.org

[9] J. Hunter, Rights Markup Extensions for the Protection of Indigenous Knowledge, 11th International World Wide Web Conference, Honolulu, Hawaii, May 2002,
http://archive.dstc.edu.au/IRM_project/paper.pdf

[10] New Methods in the Care of American Indian Artifacts, K. Dunman, Fall 2001,
http://www.focusanthro.org/american_indian_artifacts.htm

[11] Dublin Core Metadata Initiative, http://dublincore.org

[12] CIDOC Conceptual Reference Model, http://cidoc.ics.forth.gr

[13] XML Schema Language, W3C Recommendation, 02 May 2001,
http://www.w3.org/XML/Schema

[14] Java Database Connectivity(JDBC), http://java.sun.com/products/jdbc/

[15] XML Encryption Syntax and Processing, W3C Recommendation, 04 March 2002,
http://www.w3c.org/xmlenc-core

[16] XML Digital Syntax and Processing, W3C Recommendation, 14 February 2002,
http://www.w3c.org/xmldsig-core

[17] Security Assertions Markup Language(SAML), http://www.saml.org/

[18] Secure Socket Layer (SSL), http://wp.netscape.com/eng/ssl3/

[19] Ara Irititja Archival Project, South Australia, http://waru.org/arairititja

[20] Robert Sullivan, Indigenous Cultural and Intellectual Property Rights - A Digital Library Context, D-Lib Magazine May 2002 Volume 8 Number 5

[21] Digital Collectives in Indigenous Cultures and Communities Meeting, Hilo, Hawaii, August 10-12, 2001, http://si.umich.edu/pep/dc/meeting/meeting.htm

[22] Special Issue on Digital Technology and Indigenous Communities, D-Lib Magazine, March 2002, http://www.dlib.org/dlib/march02/03contents.html