Preserving Endangered Languages using a Layered Web-based Archive

L. Poulo, J. Paihama, M. Mohamed Nour, H. Suleman
Department of Computer Science
University of Cape Town
South Africa
{lpoulo,jpaihama,mnour,hussein}@cs.uct.ac.za

Abstract:

Many human languages, an essential part of culture, are in danger of extinction. UNESCO estimates that at least a half of the world's 6500 spoken languages will disappear within the next 100 years. This problem can be addressed to some extent by computer systems that collect, archive and disseminate dictionaries for various languages, thus performing the key function of preservation.

The approach taken in this project was to develop a Web-based multilingual thesaurus, with mechanisms for the submission and retrieval of language data and metadata. This thesaurus was built on top of the FEDORA Web-based digital repository toolkit. Two distinct user interfaces were then developed as part of a proof of concept language preservation system, namely a Web interface and a cell phone interface. These were created using AJAX and J2ME+GPRS respectively.

Both user interfaces were designed using an iterative User-Centred Design approach, and the back-end system was designed to meet the needs of the user interfaces, with a Web-based API.

The resulting system proved to be useful as users indicated that they could preserve spoken languages by submitting and retrieving words in their own languages. The independent successful evaluations of the 2 user interfaces together demonstrate the feasibility of creating a preservation-directed archive as a layered Web-based digital repository, where the preservation function is separable and accessible through a well-defined Web-based API.

Keywords: Language Preservation, Digital Repository, User-Centred Design, User Interface

1. Introduction

Many languages are in danger of being lost and, if nothing is done to prevent it, it is estimated that half of the world's approximately 6500 languages will disappear in the next 100 years [1]. Total death of a language occurs when a language that was previously used by a certain community is no longer spoken. This brings about the loss of cultural heritage, for language is a unique medium for its traditions and culture. Language data are central to the interests of social science research communities, including linguists, anthropologists, archaeologists, historians, sociologists and political scientists interested in culture of indigenous people. Besides popular and official world languages, many people are fluent in a diversity of regional dialects. Over time many of these languages fade away and with this significant elements of culture and history are lost.

It may be possible to address this using a computer system to collect, archive and disseminate dictionaries for various languages. The system should allow for searching and browsing through dictionaries, perform translations from one language to another, and include etymology and annotations.

WordBank was developed in response to the problem of dying languages. WordBank is a component- and Web-based multilingual thesaurus with a service-oriented architecture and a mechanism for the submission and retrieval of language data and metadata. WordBank comprises two distinct interfaces, namely: an AJAX-based Web interface and a J2ME-based Cellphone interface from which the users could submit and retrieve language data; and a back-end archive to store data and metadata which will allow retrieval of information on different languages from the archive via the interfaces. The back-end archive was built on top of the Flexible and Extensible Digital Object and Repository Architecture (FEDORA) [5] open source digital repository system.

For the purpose of this study, the chosen languages were Arabic, Portuguese and Sesotho, primarily because they are the researchers' native languages. However, WordBank can archive any other language.

The rest of this paper is organised as follows: section 2 is about related systems, section 3 is about the implementation details, section 4 is about the user evaluation conducted and finally section 5 concludes the paper.

2. Related Work

2.1 Digital Preservation

Digital Preservation is the process of keeping digital material accessible and usable for a defined period of time [10]. Digital material [3] refers to any material which can be processed by a computer and includes both born-digital resources and digitised (put into a digital form) resources. Digital preservation starts with a sequence of organized tasks and technical strategies for ensuring that the digital material is stored appropriately [3]. This, combined with adequate maintenance, can maintain availability of digital material for a lifespan of decades or even centuries.

To ensure the preservation of digital information it is very important to choose the appropriate storage medium, for example CD-ROM or DVD [17]. If the digital medium becomes obsolete before the information has been copied onto another medium, the data will be lost. The entity responsible for the data preservation must take into consideration the lifespan of the medium used for storage and refresh the data in time to prevent any loss. This can be achieved by making use of either migration or emulation.

- Migration is the conversion of a digital object to a format that is independent of the particular hardware and software that was used to create them. The main objective of migration is to ensure that an old digital object can be accessed using new technologies [17]. The Open Archival Information System (OAIS) model [18], breaks migration into four categories: refreshment, replication, repackaging and transformation.
    - o Refreshment guarantees that an accessible copy of the digital object is preserved.
    - o Replication and repackaging ensure that a manageable package of the digital object is available.

    - Transformation modifies the digital object.
- Emulation is designing software and/or hardware that will mimic the behaviour of an operating system. This will preserve the performance and the look and feel of the digital object [17]. While an emulator mimics the behaviour of old hardware platforms and operating system software, it does not involve preserving the old hardware and software.

Digital preservation is not restricted to born digital objects. We now see to a greater extent systems that preserve data and metadata about physical objects. Digital preservation is not restricted to the use of digital libraries; it is also done using other techniques/technologies. Some recent preservation projects are:

- The Nautical Archaeology Digital Library (NADL) [19] assists nautical archaeologists in the construction of ancient ships and the study of shipbuilding techniques. This project catalogues, stores, and manages artifacts and ship remains along with its associated information produced by an underwater archaeological excavation as well as manuscripts that date back to the 16th century. The system provides a visualisation tool that help researchers manipulate and analyse artifacts and their relationships as well as algorithms and visualization mechanisms for ship reconstruction (for example to help the archaeologists determine how to reassemble a ship from the fragments recovered).
- The World Digital Library [20] launched by UNESCO in April 2009 is a free multilingual Web portal that allows access to important resources from several cultures around the world, including manuscripts, maps, rare books, musical scores, recordings, films, prints, photographs, architectural drawings, and much more. The objectives of the World Digital Library are: to promote international and inter-cultural understanding and awareness; provide resources to educators; expand non-English and non-Western content on the Internet, and contribute to scholarly research.
- The Ithaka/Aluka Mellon initiative includes a database of African cultural heritage and landscape [22]. Many of Africa's architectural heritage sites are in a bad conservation state and deteriorating. The aim of this project is to provide spatial and content data as well as the creation of a permanent digital record of historical architectural sites in Africa for future reference for students and academics. The data also can be used for restoration and reconstruction purposes. The spatial database content is created by using: laser scanning, photogrammetry, remote sensing, GIS, databases and visualization techniques.
- The National Foundation of Scientific Computation (FCCN) in Portugal [23] is working on a project called The Portuguese Web archive (AWP). This project started in January 2008 and the main objective is to preserve the information published under the .pt domain. In addition to contributing to the preservation of historic and cultural digital information the AWP hopes to: increase the use of Portuguese as a language for communication on the Web; provide resources to scientific research communities in Portugal; and reduce the national dependency on international information resources. The system collects information from the Web and saves it in ARC format. The information is them replicated and saved and maintained in different locations to avoid permanent loss in case one of the servers crashes. The information archived is automatically classified by topic.
- The History Makers Digital Library [24] is a project of the History Makers organisation, a non-profit institution that records, preserves and disseminates video interviews that unveil the accomplishments of African-American groups. They aim to provide resources for exploring the African-American history and culture. This digital library came to life in 2007 as a result of collaboration between the History

Makers and Carnegie Mellon University (CMU). Speech alignment, image processing and language understanding technologies allow multiple levels of access and viewing of the videos in a large oral history corpus. To date they have created an archive of 400 interviews (1200 hours of searchable video) [25].

2.2 Language Preservation Systems

Most of the above digital library systems are used for cultural heritage preservation, but not necessarily the preservation of human (or spoken) languages. They store facts relating to a certain heritage but not language data itself.
Some language preservation projects are:

- A Wayne State University project created a distributed digital library for storing data of endangered languages, called Electronic Metastructure for Endangered Languages Data (E-MELD) [6]. E-MELD was supported by the need to address problems such as the lack of common standards and formats, and by the lack of supporting software, which currently impedes long-term storage, retrieval, display, and even comparative analysis of language data. In order to achieve consensus about certain aspects of archive infrastructure they focused on the three tasks stated below [6]:
    - Build a showroom of best practice for digital archives of endangered language data, where the data from *ten* endangered languages is archived in a way such as to demonstrate the best practice, and the best way to design and store material for such an archive,
    - Build a linguistic ontology which would serve as an interlingua for the various linguistic markups used, so as to allow searching of diverse material, and
    - Build *FIELD*, a tool that facilitates the work of linguists in storing endangered languages material to conform to best practice.

  The three major components of the E-MELD system are: a graphical user interface (GUI); a knowledge base (containing the ontology and query engine); and a database of endangered languages marked up in XML format [20].

- The native languages of the Americas [8] is a group dedicated to the survival of native American languages, particularly through the use of a Web portal as the interface between the organisation and its members.
- The Comanche language and cultural preservation committee [12] which proposes to change the direction of language by restoring the NU̶MU̶ TEKWAPU̶HA as a “living language” and take the language and its associated heritage into the future.
- The native literacy centre in Oaxaca: language and literacy preservation project [13] was founded by professionals and native educators to help preserve the numerous indigenous languages that have been rapidly disappearing in the central and Southern parts of America. They support natives from Oaxaca (Mexico) who want to increase their literacy skills and prevent the extinction of their language.
- UNESCO’s register of good practices in language preservation [14] aims to create a collection of positive experience reports from past and current project agents, which should be used to give the new generation of project agents a guide of do’s and don’ts in language preservation.

All the above-mentioned projects create awareness and provide guidelines or tools that strive to preserve languages that are in danger of extinction in different ways, though they rarely fall at the intersection of digital preservation and language preservation.

Below is a discussion of the technologies used for the design and implementation of WordBank.

2.3 AJAX

AJAX (Asynchronous JavaScript and XML) [2] is a set of Web application technologies that includes: standards-based presentation using XHTML and CSS; dynamic display of interaction using the document object model (DOM); data interchange and manipulation using XML and XSLT; and asynchronous data retrieval using XMLHttpRequest, with all these technologies coupled together using JavaScript.

The browser has an engine that acts as an intermediary between the client and the server. User action triggers a JavaScript call to the engine. The engine then sends an HTTP request to the server; the server processes the request and sends a response (in XML) to the AJAX engine. The purpose of this engine is to facilitate communication between the client and the server by removing the blank page waiting time [1]. When it needs data from the server the engine makes asynchronous requests for XML, all without meddling with user interaction. AJAX is increasingly popular and is used in Web sites like AJAXTrans[1], Gmail[2] and Google Maps[3].

2.4 J2ME

The Cellphone interface was developed using Java Micro Edition (J2ME). The reason for choosing J2ME is portability. J2ME applications run on any Java-enabled mobile device without the need to make extensive changes to the code. Its Mobile Information Device Profile (MIDP) platform is famous (some might say notorious) as the "write once, test everywhere" standard [4][7]. MIDP provides a set of APIs that define the way cellular applications interface with the phone [4]. The Connected Limited Device Configuration (CLDC) on the other hand, is implemented on top of the operating system and it defines the Java language features and the core Java libraries of the JVM for mobile devices [4].

The mobile application communicates with the network service provider via General Packet Radio Service (GPRS) using Hypertext Transfer Protocol (HTTP) whose connections are in turn made by the J2ME virtual machine embedded in the device [4].

2.5 FEDORA

WordBank's repository is based on the FEDORA digital repository. FEDORA is an open source system for storage, management and dissemination of different types of digital objects and their relationships. The FEDORA Architecture is divided into four subsystems (management, access, security and storage) and a Web services layer [16].

---

[1] http://www.AJAXtrans.com
[2] http://www.google.com/accounts
[3] http://maps.google.com

The key features of FEDORA are [9]: support of heterogeneous data types and adaptation to new ones; the aggregation of mixed and possibly distributed data into complex objects; the ability to specify multiple content disseminations of these objects; and the ability to associate rights management schemes with these disseminations.

FEDORA's functionality [4] [16] may be broken down into a set of services as follows: repository services are for depositing, storing and accessing data; index services are ways of discovering digital objects; collection services are for joining digital objects and services into collections; naming services are for resolving then giving digital objects unique names; and finally, user interface services provide users with interfaces to access the other services. The FEDORA services are a group of Web Services built using different protocols to provide their intended service.

3. Design and Implementation of WordBank

3.1 Requirements

Wordbank's features were determined through a User Centred Design (UCD) process. Requirements for the system were collected using brainstorming sessions with users to be involved in the design process and prototyping to test user satisfaction.

Brainstorming sessions were conducted with the developers, experts and users. There was open discussion on the possible features the system could have. The features were noted and discussed in further detail.

The ideas from the brainstorming sessions were implemented in low fidelity and high fidelity prototypes by the interface designers to refine the requirements. The design process was conducted in four iterations - the requirements gathering being the first - for both user interfaces. In the second iteration paper prototypes were created and tested with users to build on the requirements gathered at the first iteration. In the third iteration the basic features of the system were implemented and the user interfaces were tested with the users. The final core user requirements are summarised as follows: submit a word; retrieve a word; word to word translation; list languages and display user information.

3.2 Design

The conceptual view of the system is shown on Figure 1.

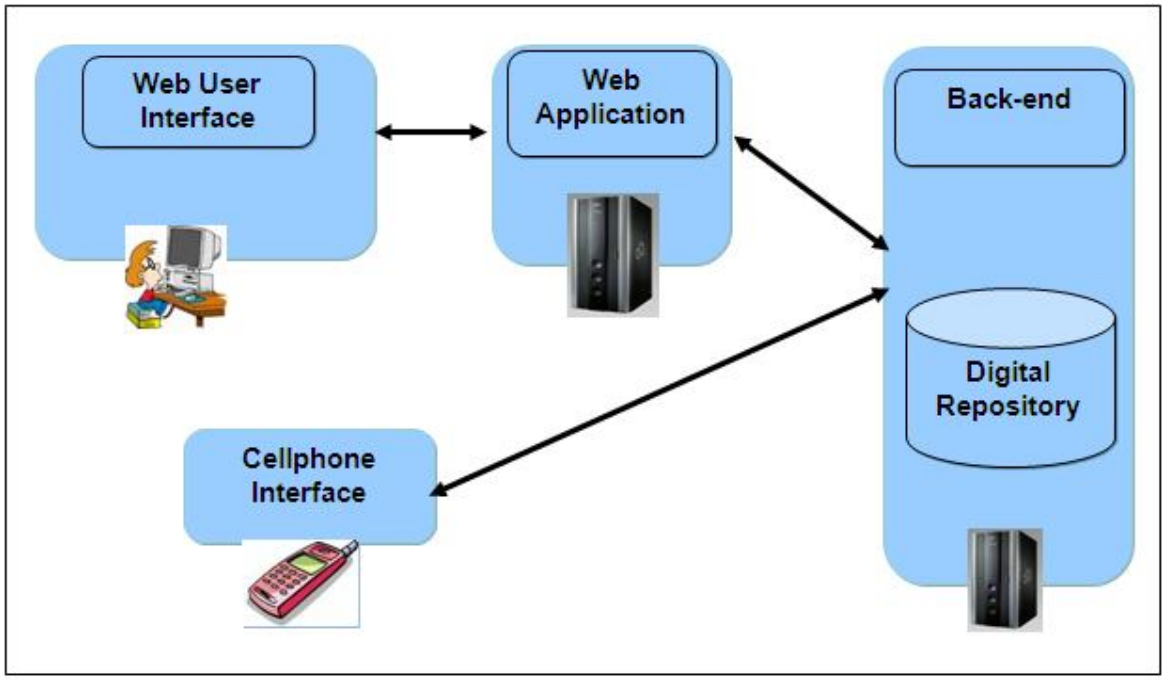

Figure 1: The conceptual view of the system

3.2.1 Web User Interface

This section presents the features of the Web user interface. The Web server and the Back-end reside on the same server as it is required by AJAX. Communication between the Web server and Back-end is established via the SOAP protocol using HTTP and XML (see section 3.3).

To facilitate navigation within the system the menu bar is fixed and appears on all pages (see Figure 2). Therefore the user can navigate as he/she sees fit and there is no mandatory sequence of button clicks the user needs to perform to navigate from one function in the interface to another. There are six main pages that can be accessed from the menu bar from anywhere in the site and then some of the pages have other child links within them.

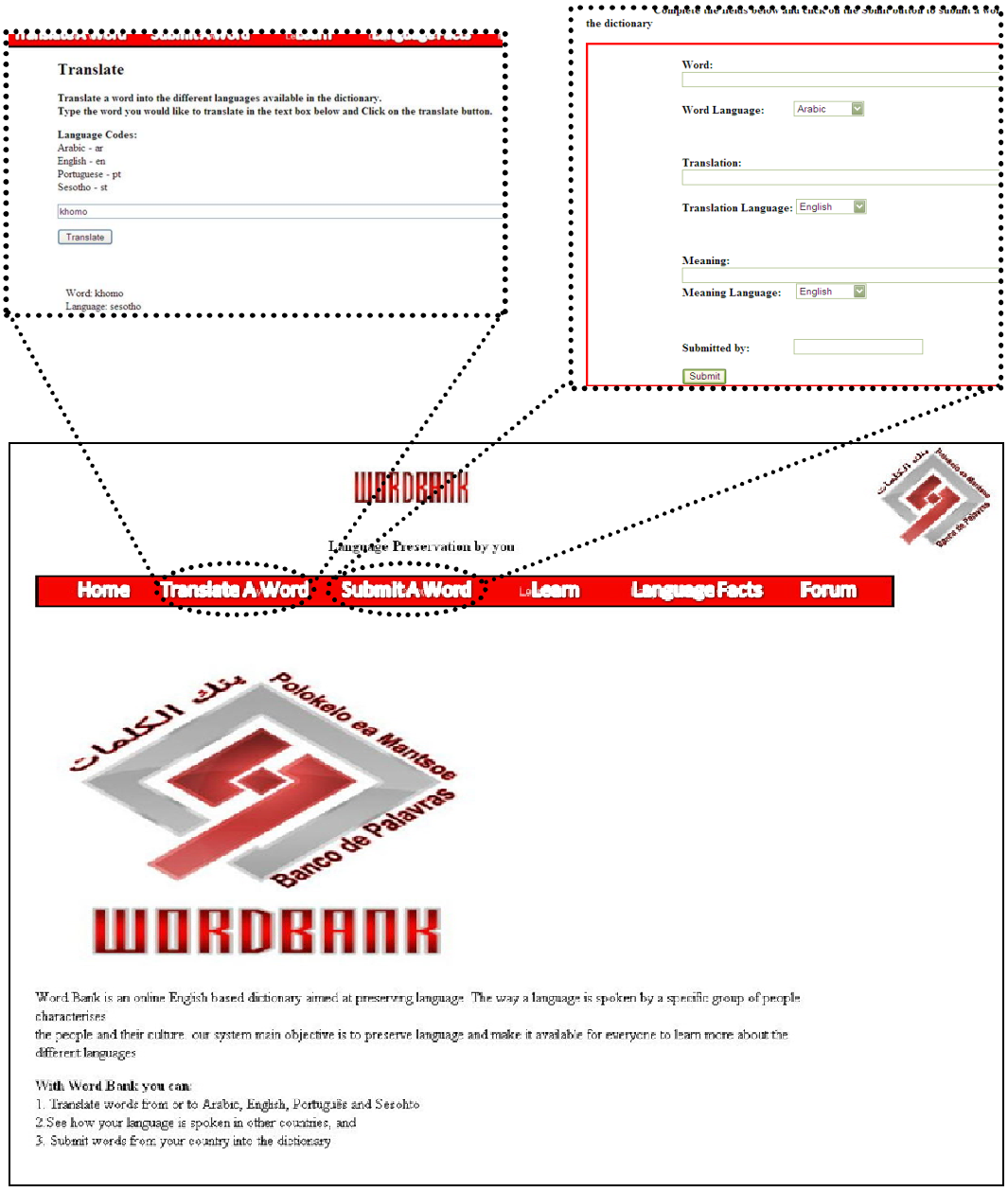

Figure 2: The Web interface

The key features of the Web interface are that it allows users to: submit a word in their language; translate a word from one language to another; learn a new language or more about their language; learn facts about a language and hold discussions in a forum as part of learning a language.

The Translate a Word page allows users to type a word in any of the four languages (Arabic, English, Portuguese and Sesotho) and then click on the translate button to get all the translations available in the dictionary.

The Submit a Word page is the page that allows users to contribute to the dictionary. A user contributes by submitting a word, its language, meaning and translation. Except for the word and the language of that word, all of the other fields are optional. This is because

the user may just want to simply preserve a word in a certain language without necessarily having to know a translation for the word.

The Learn page allows users to learn some basic concepts of the system's languages. In the Language facts page, users will be able to learn more about the background of the different languages (Arabic, Portuguese and Sesotho). Finally, forum pages allow users to get in touch with one another and post messages, comments and requests to other users.

3.2.2 Cell phone Interface

The Cellphone interface (Figure 3) is presented in this section. The key features of the Cellphone interface are the ability for users to: submit a word; perform word-to-word translation; display all languages available in the system; and allow users to view their profiles. Due to the limitations on mobile phones some of the features of the Web interface are omitted on the Cellphone interface. On a mobile phone, using HTTP over a GPRS connection, a user can explore all of the afore-mentioned features (see section 3.3 for details of the communication protocol).

From the "WordBank Services . . . " screen, the user enters a word to look for from the available dictionaries in the repository. When the user presses submit, a request is sent to the back-end from the Cellphone interface. The request is sent using HTTP over GPRS. At this stage, the word that the user has requested is checked from the returned document, and if the word is not found, an appropriate regret message is displayed to the user. Otherwise all translations of the word entered are displayed to the user.

The deposit or submit word action, on the other hand, requires that the user supplies four entries separated by spaces. If there are fewer or more than four entries, the user cannot submit the word. The four entries to be supplied are the word, word language, a translation and translation language. Input length verification happens on the front-end (that is on the Cellphone interface). If the entry is of the correct length, the request is sent to the backend from the Cellphone interface. A connection is established and a data input stream is also opened to start sending the data (words) to the repository.

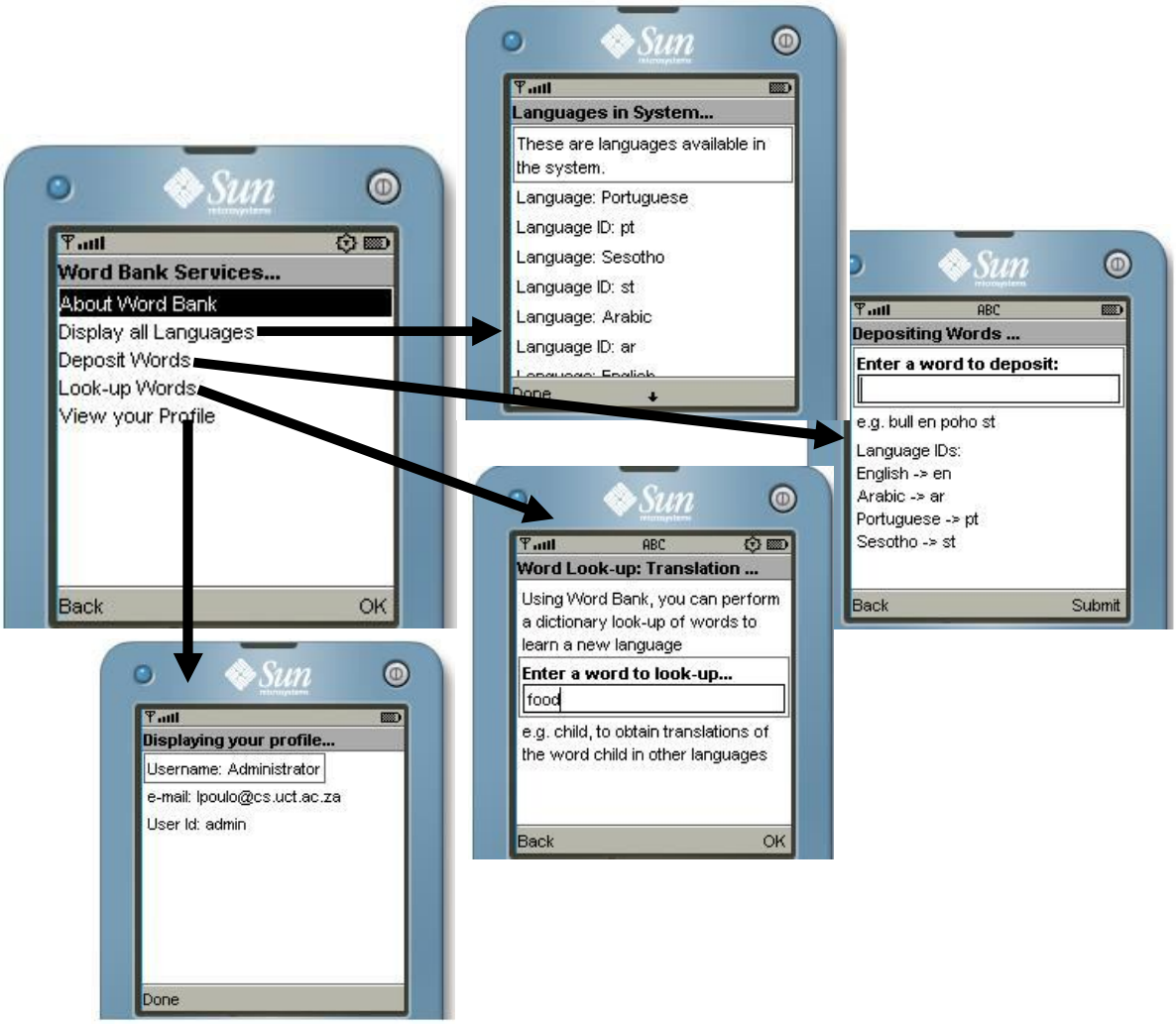

Figure 3: The Cellphone interface

3.2.3 Backend

The backend consists of a FEDORA repository and a backend application, as shown in Figure 4, which is briefly introduced in this section. Firstly, the FEDORA repository was installed and configured. Secondly, a backend application that interacts directly with the repository was implemented to allow easy interaction between the user applications and the repository. Words are stored in the repository as digital objects. A digital object in FEDORA has a PID, Datastream and Disseminator. These properties are specified by the FEDORA system: a PID to uniquely identify an object; a Datastream which is the object's content and a Disseminator that provides a way to view the object for the user. The backend application implements the features needed by both interfaces and translates these to the repository by using the FEDORA APIs. This application thus acts as middleware between these two sides of the system.

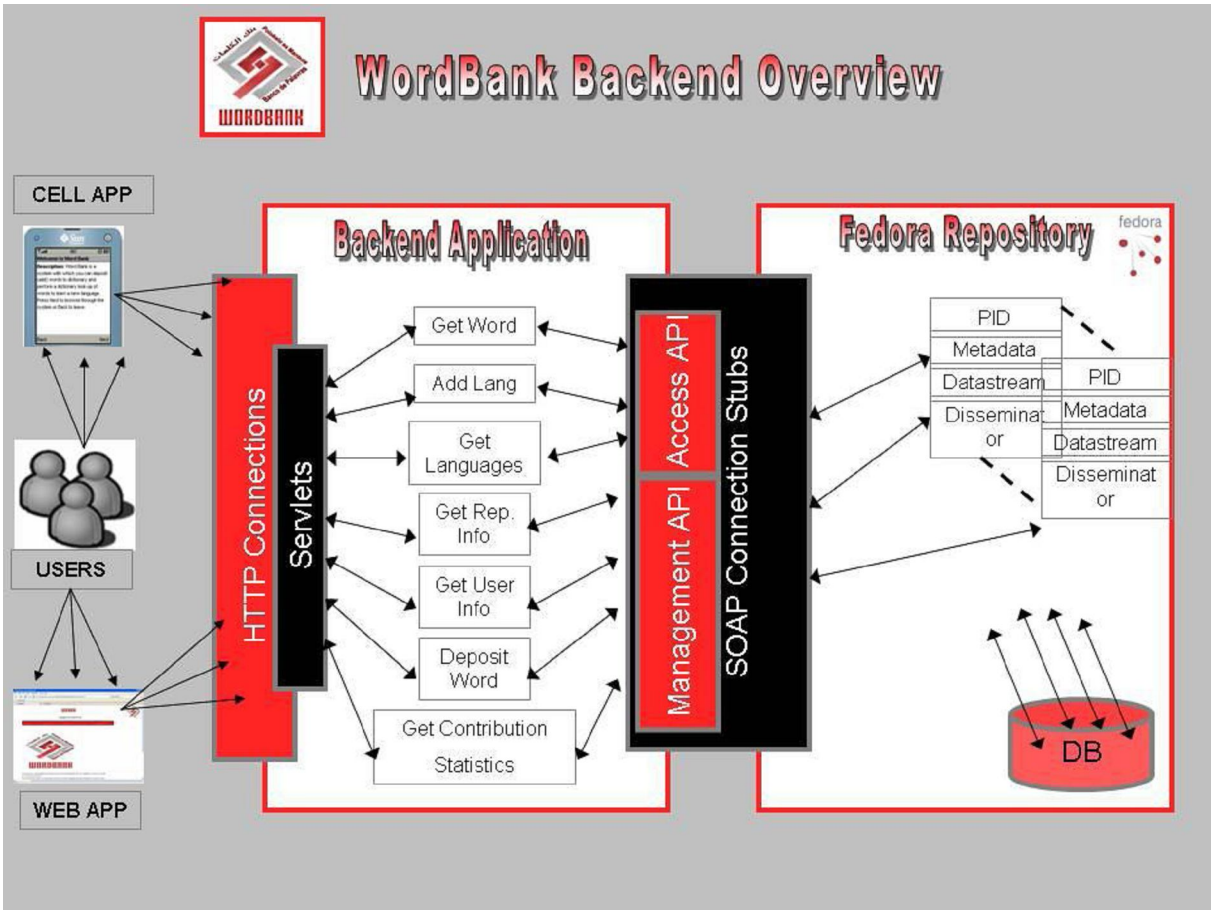

Figure 4: The back-end archive

3.3. Communication Protocols

Both interfaces communicate with the back-end via HTTP and XML. The user requests sent to the back-end are made using the HTTP GET request in order to retrieve language data from the system. Submission of language data and metadata was achieved using the HTTP POST request to the back-end. To perform word-to-word translation, the server will send back a response with all the requested fields as follows:

word-to-word translation:

Request:

http://[host]:[port]/WordBank/wb?action=getword&q=[word]

Parameters:

word

Response:

```
<entry>
    <word languageid = "id">word</word>
    <meaning languageid = "id">meaning</meaning>
    <translation languageid = "id">translation</translation>
    <user>userID</user>
</entry>
```

To submit language data and metadata to the back-end repository the application uses a POST Request method and all parameters are encoded using "application/x-www-form-urlencoding". The POST response is an XML message confirming that the word has been submitted successfully. The XML format of the POST response is shown below:

word submission:

Request:

http://[host]:[port]/WordBank/wb?action=depositword&q=[word]&wordlanguage=[wl]&user=[u]&definition=[d]&deflanguage=[dl]&translation=[tran]&tranlanguage=[tl]&eguse=[eg]

Parameters:

word,wl,u,d,dl,tran,tl,eg -
The word, word's language, username, definition, definition's language, translation, translation's language, example of use

Response:

<result>Confirmation message</result>

The application uses a GET Request to get a list of languages available in the system. The requested fields are sent as an XML file. The requested fields are language name and language id.

List languages:

Request:

http://[host]:[port]/WordBank/wb?action=languages

Parameters:

it does not take any parameters

Response:

```
<languages>
    <language>
        <name>Language Name</name>
        <id>Language ID</id>
    </language>
</languages>
```

When a user views a profile, a GET Request is used to connect to the server and the returned XML is as shown below:

User info:

Request:

http://[host]:[port]/WordBank/wb?action=userinfo&user=[u]

Parameters:

u - it takes only the username as a parameter

Response:

```
<user>
    <username>Username</username>
    <email>email address</email>
    <id>UserID</id>
</user>
```

Updating a user profile uses a POST Request method like submit word. The parameters used are email and userID. All parameters are encoded using "application/x-www-

form-urlencoding".

4. Evaluation

4.1 Method

User evaluation was conducted to determine if first-time and non-technical users could use WordBank. The methods used were direct observation and a questionnaire. The questionnaire combined scalar questions, multiple-choice questions and open ended questions that would give users a chance to provide their opinions about the different features of the system. In both scenarios the users were asked to perform exactly the same functions. First the users were given a detailed description of WordBank, and then they were given a task which they had to perform. While the users were performing the tasks they were observed. The users could ask what to do, but were discouraged from asking how, as the aim of the experiment was to check how easily they could perform the designated tasks. After completing the tasks the users had to complete the questionnaire. A list with words was provided for users who were not familiar with any of the three languages chosen for experimentation.

4.2 Results

The following were noted from the user evaluation exercise:

- On average, users spent between 5 and 15 minutes on each interface performing a pre-defined set of tasks.
- Users reported 100% satisfaction in translating and submitting a word using the Web interface and 93% (13 users) of users reported satisfaction in performing the same tasks on the Cellphone interface.
- However, five (5) users said that it was difficult to see if the request was being processed on the Web interface as AJAX does not reload the whole page. At first they struggled to see if there were changes or not, but once they saw how the interface worked, they reported that it was simple and faster than reloading the whole page. On the Cellphone interface, only one user did not understand that one had to wait for the HTTP connection to the back-end, and she thought there was no response at all.
- All users got feedback from the system, but 8 users said that it was difficult to see the feedback from the ‘Deposit’ page on the Web interface because it is hidden and they had to scroll down to see it. This led 5 of the users to keep on submitting the word because they assumed that the word was not being submitted. They suggested that the feedback be presented in a different way that is more obvious to the user.
- Out of the 14 users, 11 said it was not difficult to translate a word on the Web interface; one said it was difficult and 2 did not rate this function. On the other hand, all users found the translate word function very easy on the Cellphone interface because all they had to do was enter a word to get all the facts about it.
- Out of the 14 users, 9 said that it was not difficult to submit a word on the Web interface, 2 said it was difficult and 3 did not rate this function. On the Cellphone interface, 11 users had no difficulties submitting a word and 3 found it very confusing because they could not follow the example given on the interface.
- Users found it easy to understand and navigate through the system using both the Cellphone and the Web interface. There were mixed feelings towards the look of

the Web interface - 6 people said that they liked it, 5 said that it is uninteresting and the other 3 did not rate this question.

- Users would like to see some more help guidelines, clip arts and sound clips.
- Ten users thought that the Web interface was good the way it was, minimalistic and not cluttered, the other 4 users said that they would like to see more images. Furthermore, all users liked the simplicity and the idea of the Cellphone interface because most people are familiar with mobile phones these days.

4.3 Observations

With the Cellphone interface, some users were unhappy about the time they had to wait while the connection was being established to the back-end. Overall, the Cellphone interface was fairly easy to use and it was very trivial to predict the sequence of screens, possibly because people use mobile phones in their everyday lives. With the Web interface, users other than students were unhappy and complained that they do not understand why they should use the system. They stated that they have no interest in learning other languages because they have their own language and culture. The student users found it interesting and easy as they understood the purpose of the research and most of them understood the crucial part played by evaluation in any research.

5. Conclusion and Future Work

A multilingual thesaurus was presented together with the technologies for developing the interfaces (Web and mobile). User evaluation has proved that Web and mobile technologies can be used to build a language-independent layered Web-based archive for preservation of endangered languages. Overall 97% (100% for the Web interface + 93% for the Cellphone interface) of users were satisfied with both translate and submit a word functions.

The system's usability and functionality can be improved by adding features such as: interoperability, to allow WordBank to automatically mine data from other Web-based dictionaries; a spell-checker, to ensure the correctness of the data submitted by the users; and batch submission of words, allowing a user to submit more than one word at once to encourage users to submit more words.

Although it was a small-scale project WordBank was successfully implemented and the evaluations suggest that this approach can be used for the preservation of real endangered languages.

6. Acknowledgement

7. References

[1] Good Practice Guide for Developers of Cultural Heritage Web Services. April 2006. Website: http://www.ukoln.ac.uk/interop-focus/gpg/Preservation/#Technicalstrategies

[2] Developer Techno Lab. 2008. "*Classic Web Applications VS AJAX Web-Based Applications.*" Website: http://developertechno.blogspot.com/2008/07/classic-web-applications-vs-AJAX-based.html

[3] EMELD: Electronic Metastructure for Endangered Languages Data. http://xml.coverpages.org/emeld.html
[4] Native Languages of the Americas. Website: http://www.native-languages.org/help.htm
[5] Reference Model for an Open Archival Information System (OAIS), CCSDS 650.0-R-1, Consultative Committee for Space Data Systems, May 1999, 140 pp, http://www.ccsds.org/ documents/pdf/CCSDS- 650.0-R-1.pdf. http://www.ccsds.org/ documents/pdf/CCSDS- 650.0-R-1.pdf.
[6] Technology Reports: *Simple Object Access Protocol (SOAP)*. 2003. Website: http://XML.coverpages.org/soap.html
[7] The Comanche Language and Cultural Preservation Committee: NU͟MU͟ TEKWAPU͟HA NOMENEEKATU͟. March 2009. http://www.comanchelanguage.org/
[8] Christel, M.G, Frisch, M.H., 2008. Evaluating the Contributions of Video Representation for a Life Oral History Collection. *JCDL '08*, June 16-20, 2008, Pittsburgh, Pennsylvania, USA.
[9] The History Makers homepage http://www.thehistorymakers.com/programs/dvl/index.asp
[10] The Native Literacy Center in Oaxaca: Language and Literacy Preservation project. March 2009. http://www.literacy.org/Projects/explorer/oax_over.html
[11] The world digital library http://www.wdl.org/en/
[12] Towards Saving Africa's Architectural Heritage Sites. From our members, Vol. 10, N°4. December 2005. http://www.isprs.org/publications/highlights/highlights1205/21HL12_05FromMembers.pdf
[13] UNESCOs *Register of Good Practices in Language Preservation*. 2005. Website: http://portal.unesco.org/ci/en/ev.php-URL_ID=DO_Topic&URL_SECTION=201.html
[14] Garrett, J. J. 2005. "*AJAX: A new approach to web applications*". Website: http://www.adaptivepath.com/publications/essays/archives/000385.php
[15] Gomes, D., Miranda, J., 2008. Arquivo e Medic*ã*o da Web Portuguesa. In proceedings of CIAWI-2008. http://www.cs.cmu.edu/~christel/PDF_Papers/ChristelFrisch_JCDL2008.pdf
[16] Knyziak, T., and Winiecki, W., The New Prospects of Distributed Measurement Systems Using Java 2 Micro Edition Mobile Phone. Journal of Computer Communications, 28:pp. 183 - 193, 2005.
[17] Lagoze, C., Payette, S., Shin, E., Wilper, C., 2006, FEDORA An Architecture for Complex Objects and their Relationships, International Journal on Digital Libraries, 6(2).
[18] Lee, K., Slattery, O., Lu, R., Tang X., McCrary, V., 2002, The *State of the Art and Practice in Digital Preservation,* Journal of Research of the National Institute of Standards and Technology. Volume 107, Number 1.
[19] Lu, S. Lui, D., Fotouhi, F., Dong, M., Reynolds, R., Aristar, A., Ratliff, M., Nathan, G., Tan, J., Powell, R., Language engineering for the Semantic Web: a digital library for endangered languages, Information Research, 2004, 9(3) paper 176.
[20] Monroy, C., Furuta, R., Castro, F., 2007, Texts, Illustrations, and physical Objects: The Case of Ancient Shipbuilding Treatises. In *Proceedings of the 11th European Conference* on Research and Advanced Technology for Digital Libraries*, (ECDL)* (Budapest, Hungary, September 2007), pp 198-209
[21] Muchow, J. W., Core J2ME Technology and MIDP. Prentice Hall, 2001.
[22] Payette, S., Lagoze, C., 1998, Flexible and Extensible Digital Object and Repository Architecture (FEDORA), Research and Advanced Technology for Digital Libraries, Second European Conference, ECDL '98, Heraklion, Greece, 41-59
[23] Staples, T., Wayland, R., Payette, S., 2003, The FEDORA Project, D-Lib Magazine, 9 (4)

[24] Strodl, S., Becker, C., Neumayer, R., and Rauber, A. How to choose a digital preservation strategy: Evaluating a preservation planning procedure. In *Proceedings of the 7th ACM IEEE Joint Conference on Digital Libraries (JCDL'07)* (New York, NY, USA, 2007), ACM, pp. 29–38.
[25] Zorich, D., 2003, A Survey of Digital Cultural Heritage Initiatives and Their Sustainability Concerns, Council on Library and Information Resources, Washington D.C.