

LINKING DIGITAL LIBRARIES TO COURSES

... with Particular Application to Language Learning

Ian H. Witten, Shaoqun Wu and Xiaofeng Yu

Department of Computer Science, University of Waikato, Hamilton, New Zealand

Keywords: Digital libraries, Course management systems, Greenstone, Moodle, Language learning, FLAX.

Abstract: We describe how collections of documents built with Greenstone, an open source digital library system that is widely used internationally, can be integrated into courses offered within Moodle, a leading open source course management system. The scheme provides students with convenient searching and browsing facilities that are accessed directly from the Moodle interface. Library collections are stored on a separate server with which the course management server communicates; institutions can decide whether to run their own digital library server or use an external one. Collections are built with the standard Greenstone Librarian Interface, or with a simpler interface intended for teachers, and can be made available institution-wide or restricted to particular courses. FLAX, an extension to Greenstone, allows the text within a collection to be used as the basis of language learning exercises, under teacher control. The scheme is freely available for download as a Moodle module.

1 INTRODUCTION

Libraries have always been central to education. Yet digital libraries do not usually seem to play a central role in computer systems for educational support. While course management systems such as Moodle allow teachers to add individual readings to courses as resources for students, they do not help teachers organize collections of readings, nor do they provide integrated access to digital libraries.

In one sense, a digital library is simply a web resource that is accessed through a hyperlink just like any other resource—and course management systems certainly accommodate hyperlinks. This is an adequate solution for many purposes. For example, it allows students to access the online university library, and specialist collections such as the ACM digital library, from within a course. However, they must leave the course to visit the library. More importantly, keeping the library separate does not encourage teachers to collate their own collections of electronic readings and present them in searchable and browsable form. Nor does it capitalize on the electronic nature of the documents to allow them to be used in novel educational ways, such as raw material for language learning exercises.

This paper describes a linkage between Greenstone, an open source digital library system

that is widely used internationally, and Moodle, a leading open source course management system. The idea arose out of a project called FLAX (for *flexible language acquisition*) that extends Greenstone to automate the production and delivery of practice exercises for students who are learning English. Teachers usually want students to be able to read, search and browse the material used as the basis for exercises, and they need to do the same themselves in order to select material for the exercises they design. Furthermore, teachers often want to build their own collections of material they select, to use as the basis for exercises.

The Greenstone–Moodle link that we have built is useful not just in the context of language learning, but in any situation where teachers want to integrate customized digital library collections into their course for students to use. After discussing a general architecture for supporting e-learning with digital libraries, this paper describes the usage of a digital library collection from within a Moodle course, showing both the reader's (i.e., student's) view of the library and how teachers use standard Moodle facilities for adding collections to courses. Next we discuss how to create collections, first with Greenstone's Librarian Interface and then with an alternative lightweight mechanism designed specifically for teachers. We examine the underlying

architecture that makes this work: how the Moodle server connects to a Greenstone digital library server via a Web services interface, and the kind of communication that takes place. Finally, to illustrate how library text can benefit certain kinds of courses, we give a brief account of the language learning exercises that are enabled by the FLAX system.

2 SUPPORTING E-LEARNING WITH DIGITAL LIBRARIES

Several authors have discussed how digital libraries can support e-learning. Standard advantages of digital libraries include immense content coverage, comprehensive information retrieval services, improved accessibility through desktop access and round-the-clock availability, and so on (Lesk, 2005). Sharifabadi (2006) points out their potential in areas such as resources to support e-learning, reference services, the research-teaching nexus, and the need for instruction in library use. Dukic (2010) notes that collections tailored for particular courses can match resources to curricula, and to students' age and ability groups. Rose (2001) recommends strategies that libraries need to adopt in order to keep pace with the profound changes that are taking place in the educational system. But there are few practical proposals for linkage between course management and digital library systems—perhaps because a simple hyperlink appears to be all that is required.

In contrast, we believe that much can be gained from a more intimate connection, as sketched in Figure 1. E-learning systems fulfil at least three functions: they serve and manage course content; they manage the students in the course, including assignments, grades and so on; and they facilitate online communities centred upon courses. Our work is based upon Moodle, but the ideas are not restricted to this particular e-learning system. Digital libraries contribute to course content by providing

resources: text, audio and video—the Greenstone system works with textual and multimedia resources in a wide variety of formats. A notable feature is that end users can easily build digital library collections, so teachers can build collections of reading specifically for their courses (as Figure 1 depicts).

Digital libraries, like physical libraries, can be far more than repositories of documents. In the e-learning context they can also provide activities for students. An example is the FLAX extension to Greenstone, which works specifically in the domain of language learning. It contributes to course management by feeding back student data such as grades. It provides group activities, both competitive and cooperative, that contribute to the online classroom community. And it includes a chat server for student-student communication.

3 ACCESSING A DIGITAL LIBRARY FROM MOODLE

In our scheme, a digital library collection appears as a hyperlink inside a Moodle course. Clicking it takes you to the collection itself, embedded within a frame on the standard Moodle page. Within this frame, users navigate around the library in the standard way, remaining all the while within the context provided by the Moodle frame.

Figure 2 shows an example. As the header states, this collection has been accessed from a Moodle course entitled *How to use FLAX*. The digital library collection itself is called *The Best of Password*, and the illustration shows Greenstone's standard *About this Collection* page. The *Search*, *Browse by Title*, *Browse by Level*, and *Activities* buttons provide access to the collection's search and browsing facilities. Below these buttons is some collection-level metadata, produced by the collection designer that summarizes the collection's origin and contents.

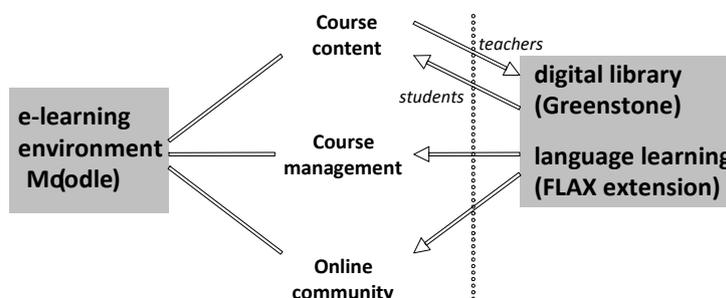


Figure 1: Connecting an e-learning environment to a digital library.

3.1 Using Greenstone Collections

A key feature of the Greenstone digital library software is that search, browsing and presentation facilities can be radically different for each collection (Witten *et al.*, 2010). They are designed and incorporated into a collection by the person who creates it—typically a librarian—and depend on the metadata that is available.

The example of Figure 2 shows a search button and three browsing buttons. The first button provides full-text search, which in this collection simply allows users to search the complete document text. The *Browse by Title* button generates the screen shown on the left of Figure 3, giving an alphabetical list of titles in the collection—which in this case is rather small, with just 20 documents. *Browse by Level* generates the screen on the right of Figure 3, which allows browsing by “difficulty level” metadata—in this case it takes on four values, levels 1 through 4. (The final button, *Activities*, relates to the language learning activities that the collection provides, and would not appear in collections not designed for this specialist purpose.) Figure 4 shows a document in the collection, reached by clicking an item in the browsing panels in Figure 3 (or the search results).

Greenstone provides comprehensive facilities for customizing collections: the searching and browsing facilities, the final document display, and the presentation graphics. What is illustrated here is just

one example. When searching, it is possible to search not just the full text but any textual metadata; additionally, there are options for stemming, case-folding, and form-based Boolean search. The browsing facilities illustrated in Figure 3 are called “classifiers,” and can accomplish far more than is illustrated here. For large collections, the document list is preceded by an alphabetic A–Z selector (this is automatically suppressed for small collections like our example). Classifiers can be based on any metadata. They can have hierarchical structures like the two-level mechanism shown on the right of Figure 3, but extended to arbitrary numbers of levels. They can be based on date, in which case the breakdown is by year and month. The information shown is not restricted to document title (as in Figure 3) but can include any metadata, formatted in arbitrary ways.

The final document display shown in Figure 4 is simple HTML, but Greenstone works with scores of different document types (including PDF and Microsoft Office documents) and can display these forms instead (or as well). The display mechanism accommodates metadata-only “documents” like library catalogues (in many different formats), and multimedia “documents” like MP3 audio files. All web pages that Greenstone generates can be configured to display their information in arbitrary ways—and in different languages. Interfaces to Greenstone exist in more than 50 languages, from Arabic to Vietnamese.



Figure 2: A Greenstone collection, accessed from Moodle.

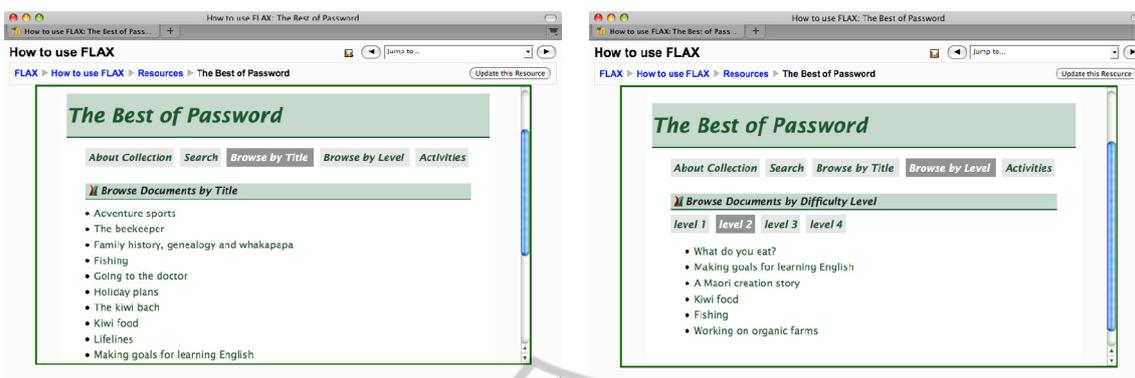


Figure 3: Browsing by title and by difficulty level.

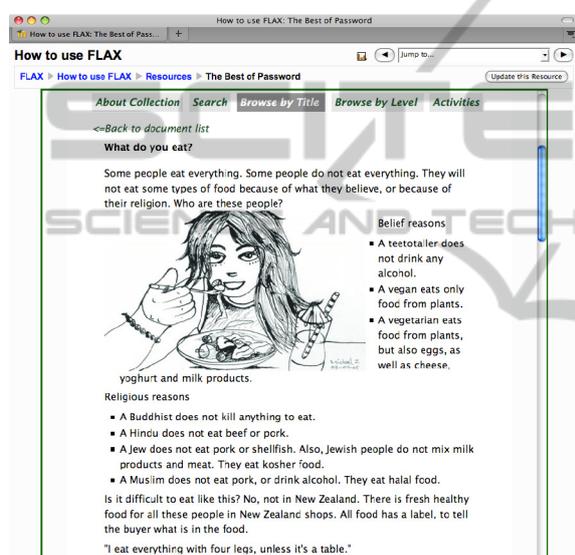


Figure 4: Reading a document.

Mousing over an entry in this list gives the brief description shown in the collection’s *About* page (Figure 1). Collections that the current user has created can be moved from the “available to your students only” category to the “shared within this institution” category, and also to an “under construction” category if the user wishes to alter the collection—for example, by adding new documents (Section 4.2).

4 CREATING DIGITAL LIBRARY COLLECTIONS

The Greenstone digital library system provides a standard way of building collections, called the Librarian Interface and designed for use by professional librarians. In addition, an alternative, lightweight, mechanism has been incorporated into the Greenstone–Moodle server, which is designed for use by teachers who want to build simple collections quickly.

4.1 Greenstone Librarian Interface

The Librarian Interface is designed for creating and maintaining digital library collections. It helps librarians and others who compile electronic anthologies to construct and organize collections of information. Only a few minutes are needed to set up a collection based on a standard design, assuming that documents and metadata are already available in electronic form.

Actually building the full-text indexes and browsing structures that support a collection may take longer. Some collections contain gigabytes of text; millions of documents. There may be even larger volumes of associated information—typically

3.2 Adding Collections to Courses

Greenstone digital library collections are added to Moodle courses in the same way as other resources. Figure 5 shows the operations involved. The *Add a resource* menu (top left) contains an item *Link to a digital library collection* that takes the user to the standard *Adding a new resource* panel (lower left), which is used for naming the resource and providing a descriptive summary.

A button at the bottom of this panel takes the user to the panel shown on the right of Figure 5 entitled *Link to a digital library collection*. (A second button allows users to create new collections, described in Section 4.2.) Here, collections are listed under three categories: ones distributed with the FLAX system, ones that are shared within the institution that runs this Moodle site, and ones that are available only in the current user’s courses.

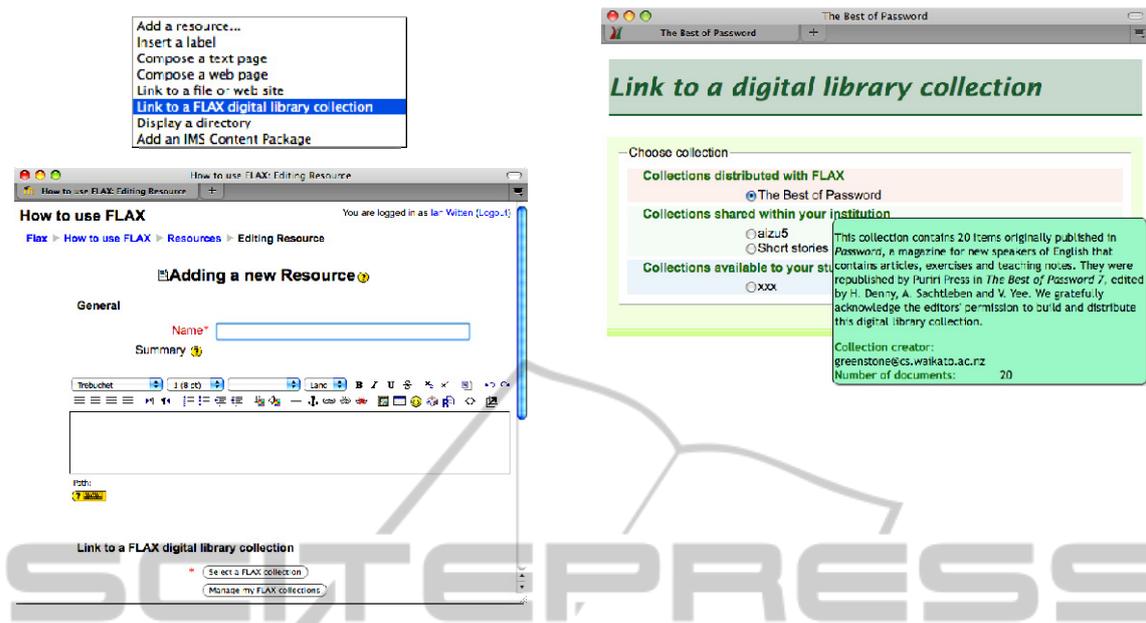


Figure 5: Adding a digital library collection to a course.

audio, image, and video, perhaps with textual metadata. Once initiated, the mechanical process of collection building may take a few moments for a small collection or several hours for a multi-gigabyte one that involves many full-text indexes. The Librarian Interface monitors all this and provides visual feedback of progress.

Naturally, customized collections that have their own idiosyncratic requirements—as most substantial collections do—take longer to set up, and the design and debugging process can take days; weeks if iterative usability testing is involved. The Librarian Interface is designed to support this sort of activity.

Figure 6 shows two example screenshots. The Librarian Interface has six different panels that are selected from the bar at the top. On the left of Figure 6 is the *Gather* panel; here users drag files and folders from their computer, represented by the blue area on the left, into their nascent collection, which is the green area on the right. The right-hand half of the Figure shows the *Create* panel that is used to build the collection and install it into the Greenstone server. Other panels allow users to add metadata to the documents and configure the full-text search and browsing facilities that the collection offers.

4.2 Lightweight Collection Building

The Librarian Interface is widely used throughout the world for building digital libraries (Witten and Bainbridge, 2007). It has a great many advanced capabilities and is necessarily rather a complex tool.

In our work on digital libraries for language learning, we have found that teachers often want to build their own simple collections of material, but lack the background and skills to work with the full Librarian Interface. Also, doing so involves leaving the familiar Moodle environment, which many teachers are reluctant to do.

For this reason we have created a far simpler interface for building rudimentary collections that can be accessed directly from Moodle by clicking the last button on the *Add a new resource* panel shown on the lower left of Figure 5, called *Manage my collections*. This takes you to the screen shown at the top of Figure 7, which gives all the collections built by the current user. Some of the ones in Figure 4 are not shown here, because they were created by other people. And Figure 7 shows an additional area for collections that are under construction (by the current user). At the bottom left is a button labelled *Create a new collection*.

When you start a new collection, it appears in the list of *Collections under construction*. Once it has been built, you can pick it up (with the mouse) and drag it into one of the other two areas, for collections that you want to share within your institution, and collections that are just for your own students respectively. You can also use the icons to edit your collection (i.e., add or delete documents), delete it, copy it, or preview it. However, you can only do this with your own collections, not those made by other people, and you can only edit or

delete a collection if you first drag it back into the *Collections under construction* area.

Here's how users make their own collections. Click *Create a new collection*, give the collection a name and description on the form that appears (lower left of Figure 7), and start adding documents.

Documents are added using the form at the lower right of Figure 7. Users choose the title, and, if they wish, difficulty level metadata (this is appropriate for purposes such as language learning), and then paste any text into the *Document content* box. Text can be copied from a Web page, a Word document, or anywhere else.

When finished, the user clicks *Next*, whereupon summary information appears, giving the documents in the collection and their sizes (in words and paragraphs). Once this is confirmed by clicking *OK* the building process starts. When it finishes (typically after a few minutes), the states of the new collection in the *Collections under construction* list changes so that it can be made available to students in the user's class or throughout their entire institution by dragging it to the appropriate place in the list shown at the top of Figure 7.

5 SERVER ARCHITECTURE

Greenstone digital libraries use a server that is distinct from the Moodle server. Several such servers exist; for example, we provide one that any Moodle installation can connect to for trials. When traffic from a particular institution becomes heavy, that institution should install its own digital library server (or switch to a pay-for-usage mode). Several different Moodle installations can use the same digital library server.

In order to serve Greenstone digital libraries, a special module must be installed into Moodle, called

Moodle-FLAX. A digital library server corresponds to a particular URL that is specified when this module is initialized within Moodle.

Figure 8 shows the architecture of the system. Here, two Moodle sites utilize a single digital library server. Communication is via Web service calls, and each call identifies the site in question by passing its Moodle *site_id* as a parameter. All users—students and teachers—work exclusively through Web browsers, depicted as computer terminals in Figure 8. Digital library related data is communicated to these from the appropriate Moodle server using the standard AJAX methodology. All this is arranged by the Moodle-FLAX software. As shown on the bottom right of Figure 8, the digital library server is also able to accommodate non-Moodle clients in a standalone mode, which we will not describe in this paper.

Here are some examples of the communication that takes place behind the scenes. A student user clicks a digital library resource that appears in their Moodle course and arrives at a page like that in Figure 2. This user action is processed by the Moodle-FLAX module, which creates a Web service request to the digital library server that includes the collection name, the Moodle *site_id*, and the fact that an “About” page is being requested. The Web service response includes the HTML frame that occupies most of Figure 2, above which the Moodle-FLAX module wraps the Moodle header shown at the top. Exactly the same procedure is used to request other kinds of digital library pages, such as the browsing pages in Figure 3 and the document page in Figure 4.

The Web service calls shown in Figure 8 are able to specify several different kinds of digital-library-related information:

- requests for particular pages of a collection, as discussed above and illustrated in Figures 2–4;

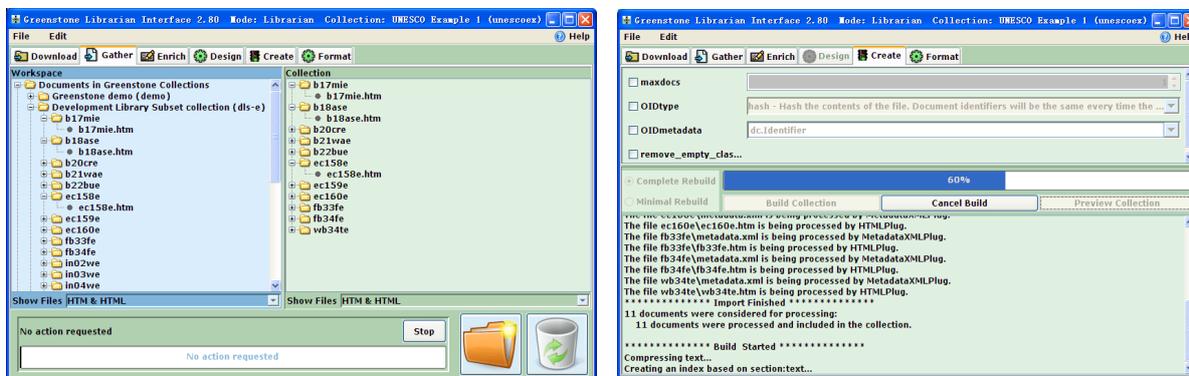


Figure 6: Building a collection with the Greenstone Librarian Interface.

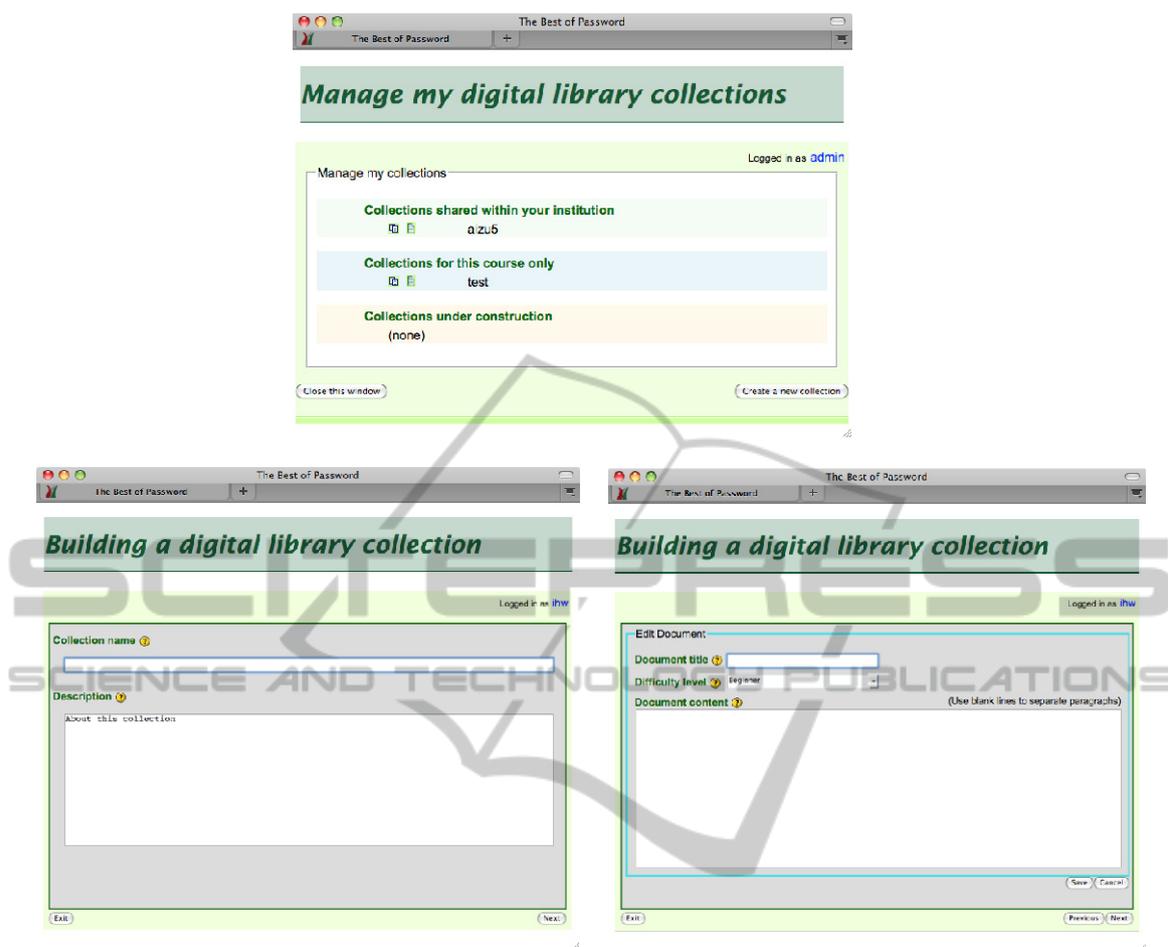


Figure 7: Managing collections, and building a new collection from within Moodle.

- a request for a list of collections for this institution and user, as shown in Figures 5 (right) and 7 (top);
- a request to change the status of a collection, e.g. to make a user's collection available institution-wide;
- requests to create a new collection and add documents to it (Figure 8);
- a request to build a collection and add it to the user's list of collections;
- a request to open a collection for editing.

Note that the Greenstone Librarian Interface (Section 4.1) operates independently of Moodle and is not accessible through the Moodle-FLAX module.

All the software described here is open source, issued under the Gnu General Public License (Gnu, 1991). The Moodle-FLAX module can be downloaded from the Moodle website, <http://moodle.org>. A demonstration digital library

server is available for public use; the necessary software for setting up one's own digital library server can be downloaded from <http://flax.nzdl.org>.

6 APPLICATION TO LANGUAGE LEARNING

A novel use of digital libraries in education is in language learning. Digital libraries include an unprecedented supply of authentic linguistic material in the form of top-quality prose. They make linguistic material easily accessible through purposeful searching and browsing. Some include rich metadata that can support interesting linguistic exercises. They provide a safe and controlled learning environment. Socially oriented software can allow collaborative activities that strengthen and enrich the students' learning experience. Finally, by building their own collections as described in

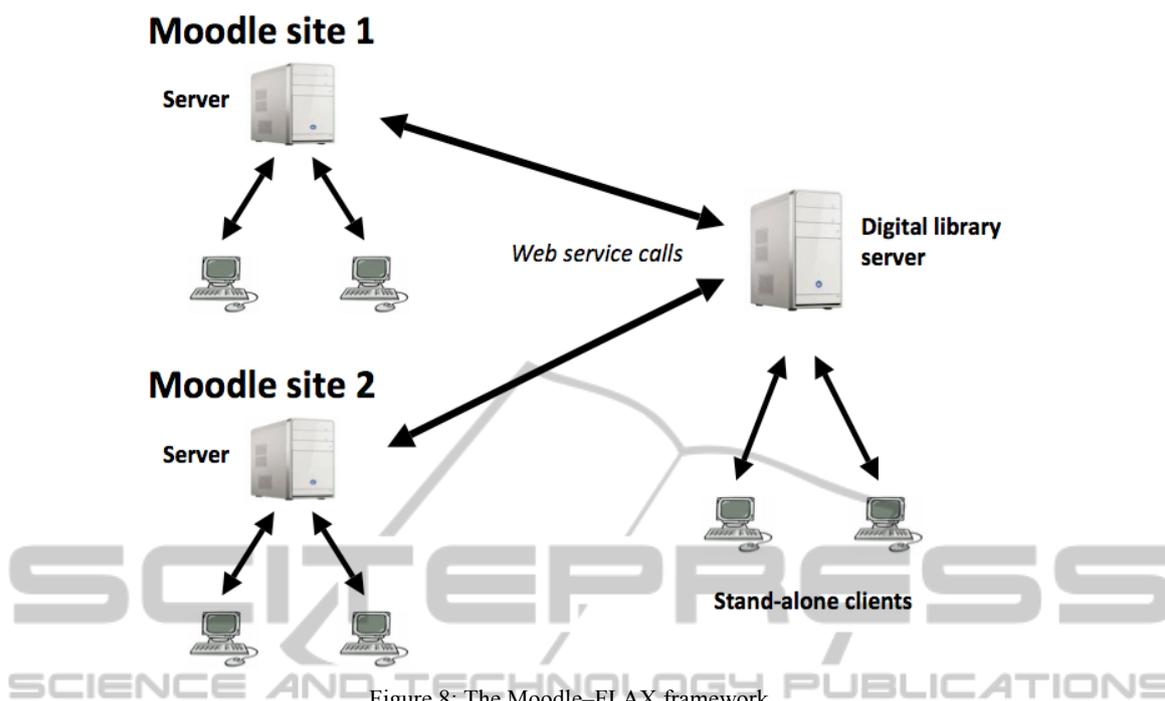


Figure 8: The Moodle-FLAX framework.

Section 4, teachers can, if they wish, utilize favourite existing material or focus exercises on a particular subject. For example, student knowledge of business language can be greatly enriched by basing learning on a corpus of business reports and product reviews (Fuentes, 2003).

The FLAX software allows teachers to create exercises drawing on digital library content that they can add to their Moodle course for students to practice. The Moodle-FLAX module integrates with Moodle's grading system, so that exercises can be assessed if the teacher so desires.

Figure 9 shows four language learning exercises. In the first, *Word Guessing*, students enter words into the gaps, based on the context within a given article, individually or collaboratively. In the second, *Scrambled Sentences*, the words of sentences are scrambled and students must sort them into their original order. In the third, *Image Guessing*, a randomly chosen image is shown to one player (called the "describer"), while the other player (the "guesser") must identify it by asking questions. Finally, in *Predicting Words and Phrases* students collaborate to predict words they think will occur in a given text.

The most striking aspect of these exercises is that they can be easily created by teachers from material that is automatically drawn from a digital library. Language exercises are added using Moodle's standard *Add an activity* dialog shown at the top of

Figure 10, which, after a brief interaction (not shown) to select the activity type and digital library collection from which material will be drawn, leads to the Teacher's Interface at the bottom of Figure 10. This interface differs slightly for each exercise type—this example is for a *Scrambled Sentences* exercise.

The online form in Figure 10 gives teachers many options. They can select which document or documents the sentences come from—either by naming the stories or by choosing their difficulty level. They can choose between simple (single clause) or complex (multiclaue) sentences. They can restrict sentences to active or passive voice, and specify what tense they should have or what modal verb they should contain. They can specify how many words the sentences should have, and can even give a word or words that the target sentences must contain. The question-mark icons give more information about each option.

Teachers can determine how many sentences are in the exercise, the number of words to scramble, and whether the sentences should be presented in the same order or a random order that is different for each student.

Finally, teachers can review the sentences that have been chosen for the exercise and discard some if they like. The *Preview* button at the bottom of the form gives the sentences proposed for the exercise. If some are unsuitable, a *Discard* button discards.

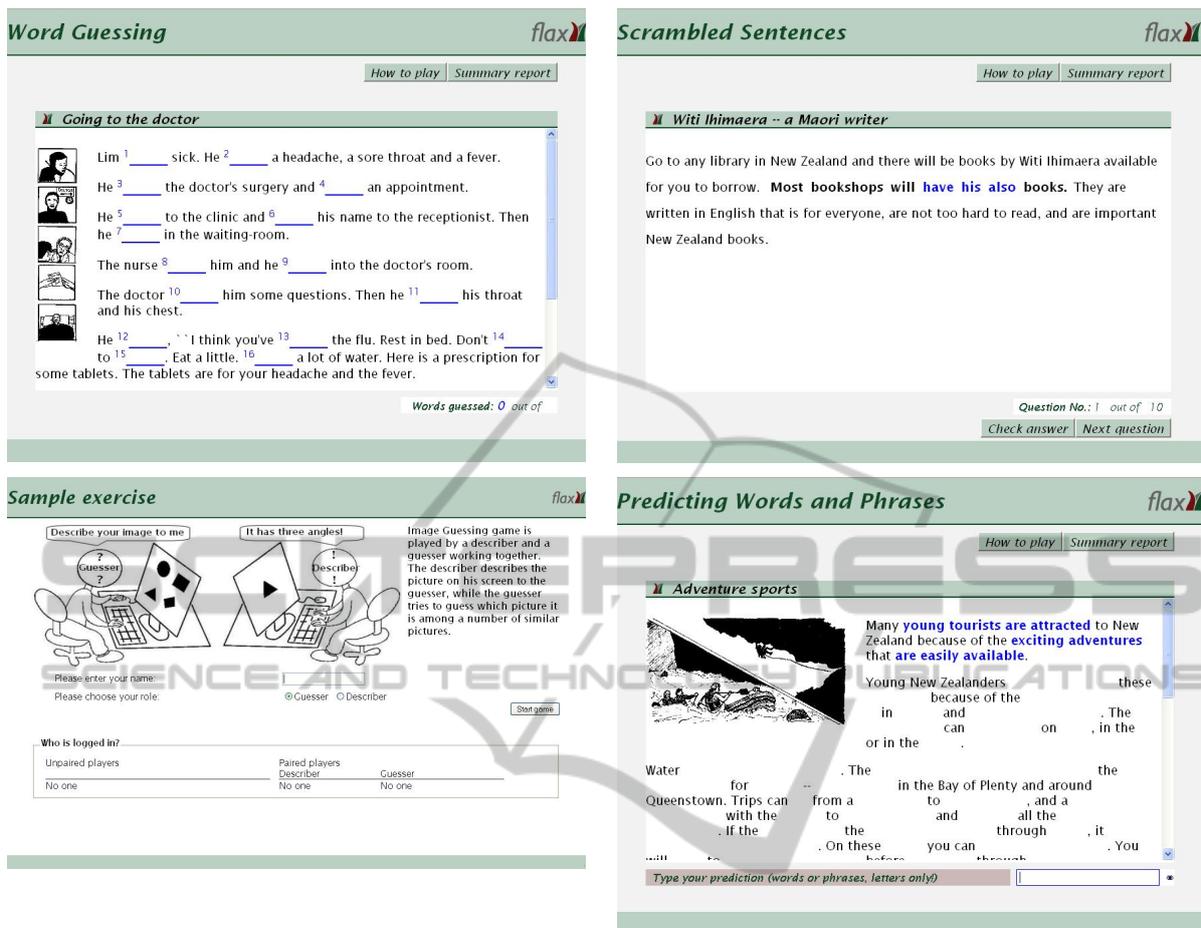


Figure 9: Four FLAX language learning exercises.

An offline version of the exercise can be generated using the *Print* button.

The FLAX language-learning environment provides many other features. Some involve all students in the class, and operate in real time—the necessary Web services to support this are built into the FLAX server. For example, exercises can take place in a competitive classroom environment, with a scoreboard displayed on each student’s screen that is continually updated with results from the whole class. Some exercises, for example *Image Guessing*, are collaborative: they are “information gap” tasks that require students to communicate, which helps them improve their communication skills and vocabulary (Pica *et al.*, 2006). FLAX includes a real-time chat facility for this purpose.

7 CONCLUSIONS

Support for digital libraries is, in general, neglected by course management systems. In one sense this is understandable: courses can provide a simple link and students can open a new browser tab or window so that they can access the library without losing their window onto the course.

However, there are advantages in providing an integrated connection between course and library:

- students remain within the context of the course while using the library;
- teachers can create their own digital library collections of course readings;
- they can do so directly, from within the course management software;
- their collections can be shared between courses and within the institution (or not).

These may seem minor, but are likely to have a noticeable effect in practice. Even more significant, in the area of language learning, is that the material in digital library collections can be used as the basis for language exercises, under teacher control—and student performance can be fed back to the grading system. This gives a new dimension to language learning.

research is funded by the New Zealand Foundation for Research, Science and Technology.

REFERENCES

Dukic, D. (2010) “Integrating a digital library into the school e-learning environment.” *CITE Research Symposium*, Hong Kong.

Fuentes, C. A. (2003). “The use of corpora and IT in a comparative evaluation approach to oral business English.” *ReCALL*, Vol. 15, No. 2, pp. 189–201.

GNU (1991) *General Public License*. Free Software Foundation, Inc.: Boston MA.

Lesk, M. (2005) *Understanding digital libraries*. Elsevier.

Pica, T., Kang, H. S., and Sauro, S. (2006) “Information gap tasks.” *Studies in Second Language Acquisition*, Vol. 28, pp. 301–338.

Rose, H. (2001) “Digital libraries and education: Trends and opportunities.” *D-lib Magazine*, Vol. 7, No. 7/8.

Sharifabadi, S. R. (2006) “How digital libraries can support e-learning.” *The Electronic Library*, Vol. 24, No. 3, pp. 389–401.

Witten, I. H. and Bainbridge, D. (2007) “A retrospective look at Greenstone: Lessons from the first decade.” *Proc Joint Conference of Digital Libraries*, pp. 147–156.

Witten, I. H., Bainbridge, D. and Nichols, D. M. (2010) *How to Build a Digital Library*. Morgan Kaufmann.

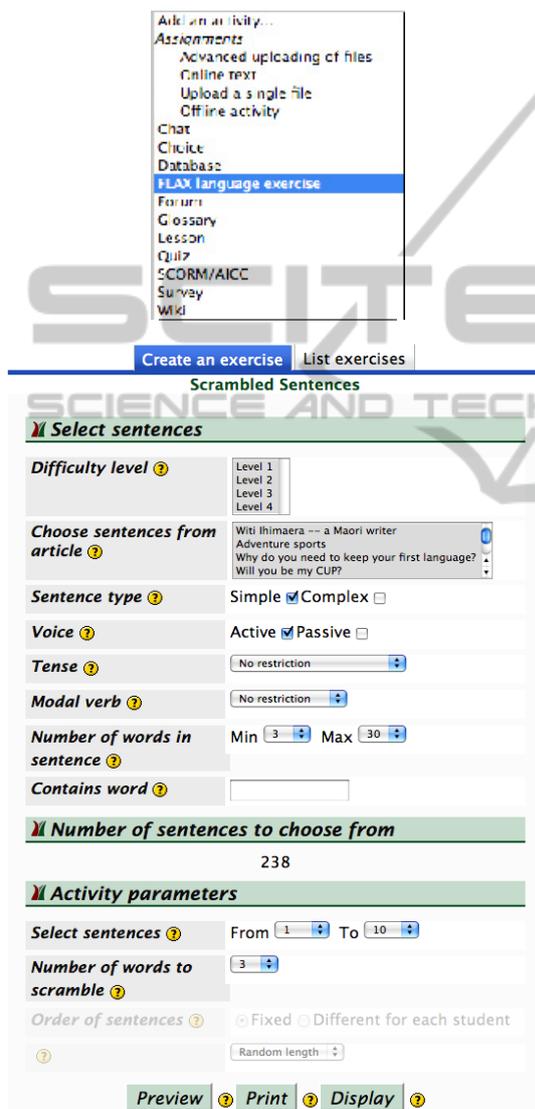


Figure 10: Teacher’s interface for *Scrambled Sentences*.

ACKNOWLEDGEMENTS

We gratefully acknowledge the stimulating environment provided by the digital library laboratory at the University of Waikato. This