Coordination of Programming-related subjects by means of an automated search engine: e-docent

Iván Álvarez, José Rafael García-Bermejo, Francisco J. García
Dpto Informática y Automátcia
inavia@usal.es, coti@usal.es, fgarcia@usal.es

Abstract – The following paper describes e-docent, a web-based application for the construction of teaching sites. E-docent is an LMS (Learning Management System). In the same vein as other similar applications, it offers various ways to ease the communication between teachers and students, and also between students, thus trying to enhance the learning experience. However, its main trait is a coordination module. This is a tool for the team of educators who are responsible for a given discipline or set of subjects. It offers support to establish a set of relationships between subjects, thus facilitating a higher degree of cohesion among related matters. This means a shift in perception for the student since he or she is provided (in an automatic and explicit way) with a set of relationships between concrete aspects of various related subjects. Students are shown a holistic view of their studies, beyond the usual teaching materials. Educators can use an automatic search system to help them find points of interest in their own material, thus selecting the most significant. Documents (PDF, DOC, PPT) supplied to the portal by the educator are supplemented with a "learning object" which contains a set of metadata that describe document contents and point to other related documents.

Keywords: Virtual environments; Web-based education; Computer Science teaching; Collaborative environments; Coordination patterns.

1 Introduction

Teaching-related institutions have seen a remarkable increase in the use of tools for the creation of education-support web sites. This includes both presential and non-presential teaching of course; a virtual campus has become of the essence. However, many educators limit their use of these sites to posting mostly static web pages, as well as various kinds of documents. This information includes lecture timetables, slides, bibliography, related links etc. Other educators go one step further and create sites in students may submit tasks, discuss between themselves in a forum, and even consult the educator. This is where LMS (Learning Management Systems) are used. Many of these solutions are open software, like Claroline (see http://www.claroline.net/) or Moodle (see http://moodle.org/). By adding modules, the applications allow the user to easily create a whole teaching model. New technologies can be integrated in the teaching model in a natural way. e-docent belongs to this set of tools.

Disciplines taught within a given curriculum usually consist of various matters, which are taught by different educators. This situation can lead to an apparent lack of inter-matter coordination, although educators are perfectly conscious of their interdependence. Thus, students lack a holistic view of their studies, and this makes them less motivated and prone to failure. The tools we have mentioned do not offer modules with an adequate support for this problem. For this reason, the Departamento de Informática y Automática de la Universidad de Salamanca, more precisely within the Ingeniería Técnica en Informática de Sistemas studies, conceived the development of a tool that might address this lack. Thus was created e-docent, a tool for the creation of teaching sites that goes a step beyond the standard modules found in other tools. Our tool supports educator coordination for the matters that belong to a given discipline or knowledge area.

From the educator's point of view, e-docent lets the user add descriptive metadata [3] to the material that is supplied for students. Through these metadata, the educator is able to establish relationships between particular aspects of his or her subject and other subjects that have been submitted to the web site. This goal is achieved by means of an automatic indexing and searching mechanism. When the educator adds a new document to the web site, the system will index it with no user intervention. Then the user can submit a set of terms that he or she thinks to be of special interest about that particular document. The system shows the user a list of documents that contain those terms; these documents may belong to his own set, or perhaps they may belong to other subjects within the site. Then the educator can set links between terms belonging to various documents that he or she considers being most significant. These links are also added as metadata for further reference.

From the student's point of view, when he or she accesses a document within the site (let's say a PDF, PPT or DOC file), e-docent will look for any metadata associated to that document and produce a set of relevant terms. They will be shown as links to other documents within the site. Thus the student no longer perceives the main terms isolated bits of knowledge, but rather as one more piece in the tapestry of his curriculum.

A first prototype has been tested on matters belonging to "Programming and data structures", which consists of four subjects: Programming, Computer Lab, Algorithmics and Data Structures. The tool has been written in Java and makes use of WebObjects (see http://www.apple.com/webobjects/), which permitted a fast development of the standard modules, thus allowing us to focus on the coordination module. The deployment platform is Apple's Mac OS X, since it provides a complete indexing and searching mechanism, Search Kit. This framework has meant a substantial reduction in the development effort.

We will now describe the various modules of e-docent, starting with the capabilities it shares with other applications and focusing then on its main strength, the coordination module. Conclusions and ideas for further work are shown at the end of the paper.

2 e-docent: a description

An essential trait of this application is the fact that it is split into modules; this was done for scalability, since we expect this tool to be used by a growing number of educators within our curriculum. Modules can be added rather easily, and the system scales up well, since for instance it is possible to spread databases, using as many machines as needed.

The Documents module is possibly the most relevant, since it relates directly to our coordination efforts. Hence we will describe it in some detail.

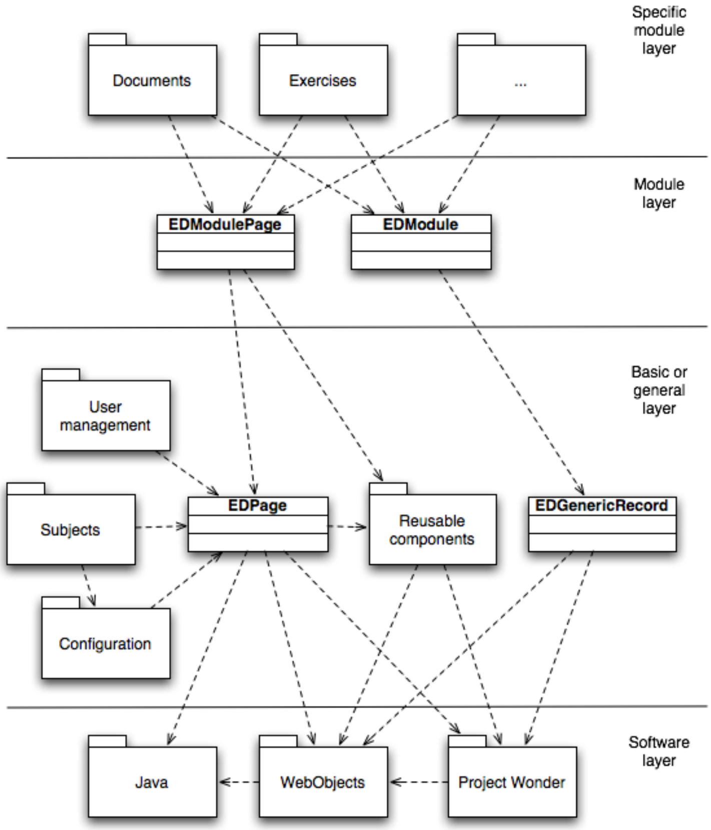


Fig. 1. System architecture

2.1 e-docent architecture

We are using a multilayer architecture (see Fig. 1). Layers are as follows:

- Basic or general layer: It takes care of site management, allowing the user to configure sites, establish common characteristics for all pages, manage users, manage subjects, create design templates, manage localization, etc. We make widespread use of reusable components, which take care of small tasks and can be found in almost all pages.

- Module layer: It permits modules to be used and included in subjects. It acts as an intermediary between the basic layer and specific modules. It carries out functions common to all modules, such as imposing access privileges.
- Specific modules: It contains the various modules that are included in subjects. Modules must comply with a set of interfaces defined in the module layer, so that they can be integrated with the system, thus sharing the functionality and benefits provided by lower layers.

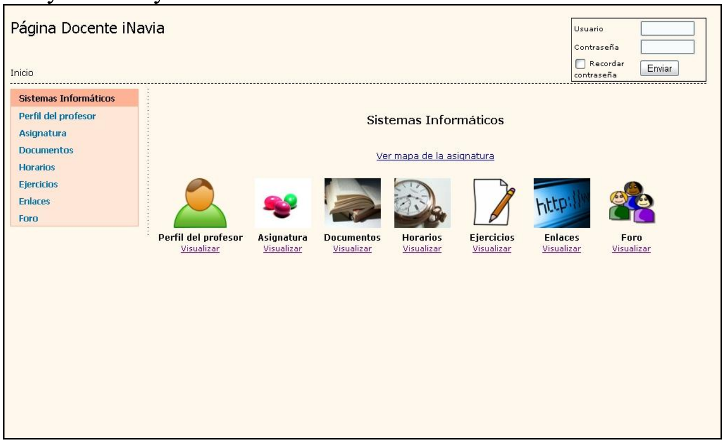


Fig. 2. Main page of an educator

2.2 Module description

Teacher information module

This module presents the user a “teacher profile” (See Fig. 2). It contains personal data, like the department he or she belongs to, his work centre, e-mail etc. A photograph may be included. This module consists of two pages, on for data entry and a second one that shows what users will see.

Subject information module

It is similar to the previous one, but its focus is general subject information. It allows the user to input data like the curriculum to which this subject belongs, its type (compulsory or optional), the year it is taught in, number of credits taught and duration. Module structure is similar to the previous one; it is purely informational and consists of a single page.

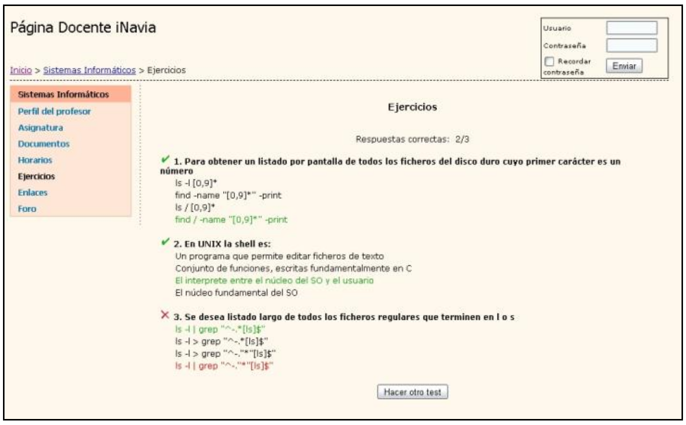
Exercise module

From the students’ point of view, this module shows an interface for self-evaluation. From the educator’s point of view, it is possible to create, edit and delete multiple-choice questions, later to generate a test based on them. Questions may have a minimum of 2 and a maximum of 6 answers, and only one of them can be correct. Questions, as well as documents, are organized in various parts or chapters within a given subject.

Questions may be tagged as public or private. Self-evaluation tests are 10 questions long; and all questions must be public. Private questions are reserved for test generation, although the educator can, at his or her discretion, include public questions in the test.

One may generate three kinds of tests:

- Fully random test: After entering the number of questions, a test is produce with randomly ordered questions.
- Structured random test: Questions are selected randomly but the chapter or part within the subject to which they belong defines their final order. This option is particularly convenient for those subjects whose parts follow a clear order, in which questions must follow that order
- Fixed-order test: The educator chooses questions one by one, and sets their position within the test.



Ejercicios

Respuestas correctas: 2/3

✓ **1. Para obtener un listado por pantalla de todos los ficheros del disco duro cuyo primer carácter es un número**

```
ls -l [0,9]*
find -name "[0,9]*" -print
ls / [0,9]*
find / -name "[0,9]*" -print
```

✓ **2. En UNIX la shell es:**

Un programa que permite editar ficheros de texto
Conjunto de funciones, escritas fundamentalmente en C
El interprete entre el núcleo del SO y el usuario
El núcleo fundamental del SO

✗ **3. Se desea listado largo de todos los ficheros regulares que terminen en l o s**

```
ls -l | grep "^-.*[ls]$"
ls -l > grep "^-.*[ls]$"
ls -l > grep "^-."*"[ls]$"
ls -l | grep "^-."*"[ls]$"
```

Fig. 3. Self-evaluation test with correct and incorrect answers

Since the process involves many possibilities, a wizard has been created to ease the creation of new tests. It conforms to the different kinds of tests chosen by the educator.

Tests are saved in a database and the educator may access them afterwards; questions may be edited or reordered.

Once a test has been created, the educator may choose to produce a preview or a printable version of the test. The pages show a test title, a header and the questions that comprise the test. The format is well suited for printing, and can be used as the basis for a written exam.

In an effort to avoid data losses when questions are erased, when one tries to erase one or more questions the systems shows a list of the exams that are making use of that question. Thus we ask for confirmation of an action that could seem unimportant, but could damage stored tests. If a question is erased that is being used in a test, this will be changed so as to close the gap, moving upwards any questions it preceded.

About self-evaluation tests, the student is shown 10 randomly chosen questions. When answers are submitted, the system shows these questions again, this time with an indication of the correct answer for each question. The total number of correct answers is shown too.

Timetable module

This module allows the educator to create weekly timetables; one can include subject-related events, such as lectures, office times and the like. Users may navigate through the weeks.

Events consist of the following data: name or event description, place, date, start and end times, and related contents. These, which are optional, are related to one of the documents supplied as teaching material. Thus an event may be created for each lecture, associating to it any contents to be used by the students.

Link module

The educator may create a list of Internet links that point to web pages with related content. Links are grouped by categories.

For each link one must provide a title, the URL of the remote object and the category to which it belongs.

Forum module

Lastly, the forum module provides an interface for managing discussion forums, in which both students and educators may participate.

Each forum consists of one or more subjects, and each subject stores messages submitted by participants. A user may send a message to start a new subject, or he or she can answer messages that belong to an existing subject.

When messages are shown, the system lays them out in such a way that users see clearly the message being answered. This is important because experience shows that some forum management programs just show a list of messages and it is difficult to see the relationship between the various messages in the subject.

Subjects in the list page are sorted by date (most recent first). The number of messages is shown also, as well as the time and date of the latest message. This lets the user see easily whether some subject may be of interest to him or her; it is also fairly easy to check for news since our last visit.

Messages store the sending user; users are allowed to edit any messages written by them if they so wish.

The educator can manage all aspects of the forum; to be precise, he or she may erase any message that seems no to be acceptable for the forum. When one is about to erase a message, there exist two possibilities: to erase the message and any answers, or to just erase the offending message. In the latter case, the message’s title becomes “Borrado” (erased) and message text becomes “Este mensaje ha sido borrado por el administrador” (this message has been deleted by the moderator)

2.3 Document module

This module is responsible for showing the educator an easy interface for uploading teaching material (documents) to the site. It must also ease coordination tasks with any other subjects in the same site. The educator may add slides, notes or any kind of content that is relevant for his or her subject. The most usual types of documents are Adobe PDF, Microsoft Word DOC and PowerPoint PPT. Documents may be added in a concrete part or chapter within the subject, or they may be assigned to a general category.

Now, more than one educator may teach any one subject. Also, any subject can be a part of more than one teaching itinerary. That is to say, students who enroll for a given subject must have passed previously some other subjects. For these reasons, an effort has been made to help the educator to establish relationships with those additional subjects.

For each uploaded document, the system saves as metadata the following information: date, title, size and subject it belongs to. The path to the document in the server is saved too. It is more efficient to save documents as files, instead of saving actual documents in a database. Thus just paths are saved in the database. It is also possible to save a set of "terms of interest" that denote concepts used in the document, but which perhaps may be present in other documents in the web site, whether in this subject or in some other(s). The systems automatically shows the educator any "teaching links" to those documents, that is to say, it shows the subject, category and particular part or chapter, besides the document pointed to. The educator may tag this link as a "related link". If so, this "related link" will become a part of the metadata associated with the document being processed.

There are two ways to add documents. One of them is the usual upload procedure, in which the educator selects a document in his or her computer and indicates any properties (title, subject etc). One other option, most convenient when one wants to upload several documents, lets the user compress (zip) those documents and upload the compressed file. It will be uncompressed in the server without user intervention.

This option is extremely convenient when starting a site, since it lets the educator zip any and all files that must be uploaded and send them over in a single step. The file will be checked for validity on the server. After checking, the system will process all of the documents stored. In this way, metadata can be added file by file in an orderly way.

We are offering the possibility to create Learning Objects following the scheme proposed in IEEE-LOM [4] for the description of teaching resources. In this way, the exported objects may be used again by other tools, or could be recovered in a new deployment of the web site without having to enter again all of the additional information. For this interchange facility we use IMS CP specification [5].

The indexing system

It is based on Apple Computer's Search KIT [1] whose main characteristics are as follows:

- It is possible to select one or more paths, depending on the documents one wants to index.

- Text extraction, getting rid of any unwanted information, such as its properties or text formatting. It eliminates any problems related to accessing information stored in documents with non-text or hard to read formats (Microsoft PowerPoint, Microsoft Word, Adobe PDF).
- One can set a list of ignored terms, that is to say, terms that will not appear in the index. This includes, for instance, articles, prepositions and also terms of a minimum length, below which they will be ignored. It is also possible to use proximity searches, which is of the essence if we want to find literal sentences.
- Support for synonymous lists. However, root extraction is not supported, although the developer may develop his or her own system (Porter Stemming Algorithm, Porter Stemmer) and can create an index of extracted roots
- Support for a minimum occurrence frequency for a given term for it to become a part of the index.
- Support for creating more than one index, index groups, etc.

Although this framework, at this time, is used basically for index generation and in order to make searches for the terms given by the educator, its architecture is designed in such a way that more complex possibilities may be added.

One other aspect of functionality at this time, which we may enhance in the future, is the student's access to these documents. Nowadays, when the student accesses a teaching material document, he or she is offered all of the information in the metadata. In this way, he or she may access links to other documents, other categories or even subjects in which these concepts are studied more deeply. Subjects no longer look like isolated islands to the student, but rather to the contrary, the student has a holistic perception of all of the information in the site, and consequently, of the subjects he or she is being taught.

On last remarkable aspect is a system by means of which the educator in charge of a given subject will be notified whenever a new document is added that contains any of the relevant terms he or she has added for his own documents. In this way, we propitiate a continued update of the metadata of any documents in store in the site, irrespective of the time when they were added.

Conclusions

This paper introduces e-docent, an LMS tool to be used as means to further the teaching and learning experience for both parties involved in the process. One of its main goals is scalability; with this purpose in mind, a multi-layer approach has been used. This has made it possible to add new modules in a simple way.

A first prototype of this tool is being tested currently. Two subjects are involved (Programming and Algorithmics), which belong the first and second semester, respectively. A small number of students have been involved in tests. The second semester is not finished yet, but many suggestions have been submitted, especially regarding usability, as well as the usual crop of bugs. A later version of the tool will take into account the indications of students and teachers, and will be tested again with a wider audience.

The most significant module is the one that deals with documents; both from the teaching point of view and because of the enhancements we are building. Beyond letting educators add files that students will then download or access, we are making this an active module. The module is in charge of searching for “related terms” inside documents stored (within the module's scope) in the site. Educators can establish “related links” with documents inside the repository, both with those produced by the educator and with those produced (possibly by others) for documents belonging to other subjects. This information (related terms and related links) becomes associated with the documents as metadata as learning objects.

In this way the system is a way to ease coordination between educators, since they can see where other educators are making use of what they teach. This is very beneficial for students, who can see that subjects are not islands but pieces in a larger tapestry. There is a more cohesive point of view for all related subjects being taught.

The indexing and searching system is reasonably simple, and we intend going further, possibly in the way of today’s web search engines, but always within the scope of our repository of learning objects.

We are conscious of the importance of standards and specifications for e-learning environments [2]. Hence we intend to make it easy to exchange, distribute and reuse our material, both in the sense of documentation and about methods, techniques and evaluation procedures.

Acknowledges

This work is partially supported by Education and Science Ministry of Spain, by FEDER funds through KEOPS Project (TSI2005-00960) and by Agencia para la Calidad del Sistema Universitario de Castilla y León funds through Portal Web, basado en un patrón de coordinación y evaluación, para las materias “Programación y Estructuras de Datos” project (US 10/04).

References

[1] Apple Computer Inc. “Search Kit Programming Guide2, (http://developer.apple.com/documentation/UserExperience/Conceptual/SearchKitConcepts/index.html), 2006.

[2] Berlanga, A., García, F. J., “Learning Technology Specifications: Semantic Objects for Adaptive Learning Environments”, International Journal of Learning Technologies. Vol. 1, Nº 4, pp. 458-472, Inderscience, 2005.

[3] Hernández, M. J., López, C., González, M., García, F. J., “Organización y Búsqueda de Contenidos Educativos: Aportaciones de la Web Semántica”, in Memorias del Encuentro Internacional de Educación Superior, Virtual Educa 2005, UNAM, 2005.

[4] IEEE. “IEEE 1484.12.1-2002 Standard for Learning Object Metadata”, (http://ltsc.ieee.org/wg12), 2002.

[5] IMS CP. “IMS Content Packaging 1.1.3”, (http://www.imsglobal.org/content/packaging/index.cfm), 2003.