Bridging the gap between teaching and breaking news: A new approach based on ESHE and ICT

Soto Montalvo a *, Jesús Palomo b, Pilar Laguna b

aDepartment of Computer Science, Universidad Rey Juan Carlos, Móstoles 28933, Spain
bDepartment of Business Administration, Universidad Rey Juan Carlos, Vicálvaro 28032, Spain

Abstract

Information overload is one of the main challenges in the current educational context. The European Space for Higher Education (ESHE) demands autonomous student’s work and that the theoretical concepts introduced in class are contextualized into the current socio-economic situation. In this paper we present iCollege, a new educational platform that compiles information from different sources and presents only relevant breaking news classified into different subjects, enriching both teachers and students learning experience. To do so, we integrate different Information and Communication Technologies, such as RSS feeds, crawling and Natural Language Processing techniques. As an alternative to the traditional practice, now the burden of scanning through a vast number of news to filter those closely related with the course is left to iCollege. The system collects, filter and classify the breaking news. A tags cloud view helps to focus on the most important up-to-dated words/events at a first sight, and eases the access to the relevant information. Furthermore, the platform, through the Wiki technology, promotes collaboration among students and provides different teaching tools to supervise the students’ job. With the current settings, every day, on average, the system retrieves more than 250 news for three courses from 44 different sources, providing an important volume of relevant information for both students and teachers.

© 2010 Published by Elsevier Ltd.

Keywords: Supervised and collaborative learning; Educacional Improvement; Automatic Generation of Knowledge; Web 2.0, Clustering; Autonomous Student Work;

1. Introduction

The rapid growth of the Internet and the wide availability of electronic documents is one of the main challenges in the information Era. This huge amount of information, given the dynamic and heterogeneous nature of the Web, makes difficult to combine it and to extract knowledge. In the educational environment, students of all educative levels use Internet as part as they learning process. The most popular news services (online newspapers, Google News, Yahoo News and Altavista News) present the information organized by related topics. However, they are too general and broad for a specific subject and, hence, the time consuming task of looking for relevant information is performed by hand. Alternatively, search engines can be used to look for a particular concept but, once again due to their general purpose, often the results are not precise –obtaining irrelevant information. Therefore, in any case, users still have to process the results obtained. Others systems that collect, analyze and display news stories, offer
extra functionalities. For example, in (Teitler, Lieberman, Panozzo, Sankaranarayanan, Samet & Sperling, 2008) users, by panning and zooming a map interface, can retrieve stories based on both topical significance and geographic region, in (Atkinson & Van der Goot, 2009) group similar articles daily, and extracts names of places, people and organizations, lists the user-defined specialist terms found, links clusters and entities, and generates hyperlinks. But once again, the information is too general for educational purposes.

The European Space for Higher Education (ESHE) raises new challenges for the traditional education. The emphasis is now put on the students' autonomous learning guided by professors, and on linking the concepts with students’ context. Following this paradigm, the traditional approach of suggesting only text books and scientific papers raise the problem of losing relevance due to the age. As an alternative, the practice of using current real examples while teaching theoretical concepts is extremely positive, since the students feel the usefulness of the topics covered in class, to understand the current sociological, cultural, economical and political environment. Given the limitation of the current general purpose information tools explained above, students and professors end up spending more time looking for useful information rather than reading it.

In this paper we propose an educational platform, based on compiled information from different sources and Information and Communication Technologies (ICT), to achieve two educational objectives: to create educative resources by topics and subjects, and to develop a new way of working oriented towards continuous evaluation. The rest of the paper is structured as follows: Section 2 presents the platform and the details of its implementation. Some performance evaluation is showed in Section 3. Finally, conclusions and future work are presented in Section 4.

2. The platform iCollege

The new platform iCollege has two main subsystems: iNews College, where students and lecturers can obtain up-to-date relevant classified information; and iWiki College, that promotes the collaborative work of students and that provides different continuous evaluation tools for the lecturer. The information flow, the educational improvement and the collaborative workflow are presented in the Figure 1.

The left hand side of Figure 1 represents the Internet with its multiple sources of information and the right hand side represents the Educational System, say faculty, students and graduates. The two subsystems are connected within iCollege, channelizing the information flow provided by iNews College (solid red lines), the educational improvement (dash-dotted black lines), and the collaborative workflow (dashed blue lines) provided by iWiki College. The first cycle involve students, following professors’ instructions, and graduates who consult the breaking news presented in the platform. The educational improvement cycle starts with the students’ active participation in class, motivated by the real examples founded in the news, followed by the comments left in iWiki College, in a critical, collective an open way. Finally, the collaborative cycle begins with the homework that will be developed in a wiki environment by each student (publicly available for the rest of the students). In this way, the quality of their works increase, and the dynamic collaboration of students is enhanced by the iWiki College.

Following the European Lifelong Learning paradigm, the tool we present is also useful for graduates, since it becomes an informative reference that can be used during their professional career -contributing to the necessary connection between the professional and the academic worlds.
The Natural Language Processing (NLP) tools are crucial in the information processing and in the generation of knowledge. We have combined these techniques with the expertise provided by professors to develop the proposed subsystems as introduced in the following subsections.

2.1. iNews College

This subsystem extracts, from online sources of information written in different languages, knowledge in an effective, flexible and automatic way, to find breaking news that is useful for a particular subject. The goal is to aid news browsing by providing an automatic, user-friendly access to classified news, and links to the original articles for further information. The system has different sequential stages through a cascade model, see Figure 2. The whole process is iterated repeatedly with a predefined frequency, say hourly or daily.

At the first stage, the system collects news by crawling reliable websites, written in Spanish and English, defined by professors. A web crawler or bot tool has been used widely for many purposes, mostly search engines. Some examples can be seen in (Best, Van der Goot, Paola, Garcia & Horby, 2002), (Dan & Mocian, 2009), (Evans & Klavans, 2003) and (Radev, Blair-Goldensohn & Zhang, 2001). In our system the starting points for the bots are the sources predefined by the lectures. In the following step it classifies them, by means of the information retrieval module, based on the syllabus of the course and the key concepts; and finally, this classified relevant information is presented in the platform. For instance, “The FOREX market” is a topic within the International Financial Markets course that has the key concepts “currency” (English) and “divisa” (Spanish), and the different sources are, e.g. http://www.economist.com or http://www.iht.com (English) and http://www.invertia.com or http://www.expansion.com/inversion (Spanish).

To determine the appropriate key concepts and sources is crucial, since both will affect the relevance of the retrieved news for a particular course. Therefore, the quality of the results of iNews College depends mainly on the criteria defined by the lecturers.

2.1.1. News Compilation

To compile the news we have a collection of RSS feeds, belonging to those sources defined by lecturers. A RSS feed usually contains a title and a summary of the news, and depending of the target application it could be necessary to obtain all the text, as occurs in (Freeman, 2009), (Li et al., 2007) and (Atkinson & Van der Goot, 2009). In our case, to guarantee a good classification we need to process all the content of the news, therefore, the stories from all these feeds are crawled.

In order to discard repeated news, it is checked first whether the website have already been visited or not. Every 12 hours the process is iterated. Since the news contents are mixed with other elements such as advertisements, images or menus, downloading news is neither simple nor trivial. In addition, the structure of the news is source dependent, so it requires an ad hoc crawler. The news contents retrieved are structured and transformed into XML format (Bray, Paoli, Sperberg-McQueen, Maler & Yergeau, 2008) for later processing.

2.1.2. News Filter and Classification

Having downloaded the news, the filter and classification stage is launched. We have developed an Information Retrieval (IR) unit to allocate relevant information in the news collection. First, we index the information from the compiled news collection in XML format. The indexing process has been done with Lucene, an open source search engine by Apache used in many NLP applications, e.g. as a support system to manage corporative information (Nurminen, Honkaranta & Karkkainen, 2005), as bilingual IR system (Zhang, Zhang & Chen, 2005), question
answering system (Jimeno & Pezik, 2007), and a management information system where Lucene is integrated to organize and retrieve documents (Passadore & Pezzuto, 2007).

The indexed content is vital in our system, since the collection of news grows on every iteration of the compilation process. Based on the key concepts, the queries are defined for each topic on every course and only the retrieved news with a degree of relevance higher than a threshold are kept for a particular topic.

2.1.3. Visualization

The students and lecturers can see the classified relevant news in a dynamic website. The most recent information is shown directly, and older news can be accessed through the historical link. Several courses are presented on the main page, subdivided in their topics. You can select the language to see the contents in the website.

For each new it is presented the headline, which is a link to the corresponding original source webpage, and the first few lines of its content. Also, the name of the source and the arrival time into iNews College, and a language flag is showed. To give a general snapshot of the current news in the platform we use a tag cloud, one per language. When a user clicks on a tag, the related information is presented in a box.

Roughly the contents might be summarized as Who does, What, to Whom, Where and When. Therefore, all the tags showed in the clouds are Named Entities (persons, organization, locations, dates, etc.). To identity and classify the Named Entities in the news we have used a Named Entity Recognition (NER) module of LingPipe, which is a suite of Java libraries for the linguistic analysis of human language. See in Figure 3 a screenshot corresponding to the course “International Financial Markets” (Mercados Financieros Internacionales in Spanish) and an example of tag cloud.

2.2. iWiki College

The synergies between iNews College and a wiki environment make iCollege a new educational platform that promotes active learning and helps both students and professors. This way, iWiki College allows the professor to supervise students’ activities, and to define activities such as: news comments, personal spaces creation, job presentations, and so on. Some advantages of the iWiki College are: simplicity in defining job deadlines; standardization of the jobs quality, because the jobs are visible by the rest of students (no changes allowed by third classmates); usability for students’ interaction; and dynamic feedback, since the students may receive comments during the activity from both professors and other students.

Figure 3. Screenshot of iNews College and the tags cloud with the main entities of the current breaking news
3. Evaluation

In order to evaluate and improve the platform, we have surveyed 145 students with different qualifications: 48 of the Master on Business Administration, 59 taking the course “Marketing Management” of a bilingual degree, and 39 taking the course “International Financial Markets” of the M.A in Business Administration. We have focused on their preferences when looking for information on the Internet, and about their perceptions when using iCollege. We found that 51.7% of the students acknowledged that they read online newspapers daily, and 70.2% in the case of the Master’ students. Moreover, 37.9% admitted the use of the results of search engines, and 18% follow the sources indicated by professors in class. This clearly shows the necessity to increase the connection between students and the real socio-economic events. As far as the iNews College is concerned, the global perception, on average, was (4), on a scale (1-5). The most relevant attributes perceived in the web site are: the relevant information for the subjects; the usability and the accessibility to the relevant information (e.g. section); the rate of updating the information, and the time gain over searching information in several sources separately. The reliability of the platform is high, because the 73.1% express that in more than 50% of times they have found relevant news for their study. Finally, the 81.4% declare that their interest for the subjects is highest after use iNews College. Moreover, the 55.2% of the asked students will use the platform once they finished the course.

We have also interviewed professors that use the platform and they agree with the usefulness of iNews College as a teaching tool due to the fact that it is time consuming to find reliable sources of information to use in class. Furthermore, the breaking news are shown in a single webpage which makes easier the navigation. Regarding iWiki College, motivated the students, increased the quality of their assignments and the professors followed their progress in an easier way.

4. Conclusions and Future Work

We have developed a new educational platform that tackles some of the challenges of the new ESHE, and that will help both professors and students.

The two subsystems in iCollege are interconnected, taking advantage of the synergies between iNews College and iWiki College. The former, among the current users, has become a reference to find knowledge related to the courses. The students easily connect theoretical concepts with the current events in a modern and interactive fashion, without spending time searching through out the Internet. The latter, iWiki College, allows the collaboration among students by learning, understanding and sharing their analysis. Furthermore, the professors can evaluate this work continuously and can propose team assignments easily.

The multilingual capability of the platform helps international visiting students (e.g. from Erasmus and/or Fulbright programs) to better integrate into the dynamic of the classes and to improve their understanding.

It can be concluded from the survey that teaching through breaking news has motivated an additional interest over the concepts introduced in class. Also, it confirms the need of this tool, not only from the lecturer viewpoint but also from the students’, since it offers important up-to-date information straightforwardly.

Although the 69% of the surveyed students consider that the number of sources of information is enough, we are developing new crawlers to access other sources of information. Furthermore, to continue facilitating the search of information in an efficient way, we will offer not only the organized information by relevance and dates, but also the possibility of making advanced searches within the platform. In addition, to extract automatically the structure of the news we could consider including a module based on machine learning.

References


