Abstract

The vision of a common international education space includes a scenario for the composition of personalised study curricula consisting of modules offered by different universities and colleges. In this scenario there is a strong need for efficient services enabling individual search and comparison of study programmes and modules. The SWAPS project, presented in this paper, is focussed on the development of a dedicated search system, based on Semantic Web technologies.

The SWAPS system facilitates semantic search for study programmes and modules. It is based on highly automated annotations of programme descriptions already available on the web.

1. Introduction

Social and cultural developments of the past three decades have led to world wide revision of the graduate and postgraduate education systems. Students have got a lot of new options for designing their personal study career.

A large variety of study programmes are available on every level of personal qualification, from undergraduate programmes to special post-doctoral and further education programmes offered for highly qualified professionals. Furthermore, study programmes have become modular, giving students the chance of more mobility during their studies. Organisational and legal efforts have been taken to guarantee high compatibility of study programmes and modules, e.g. the so called Bologna Process in Europe [1]. At the same time new instructional methods (such as blended learning) and infrastructures exploiting digital communication media and networks enable students to pursue their studies in a location independent way.

These processes of change in the education systems will go on. The student of the future will compose study modules offered by different educational institutions to a highly personalised curriculum. According to this vision the role of educational institutions, such as universities, will change fundamentally [4]. They are to be considered as components of a common, integrated education space. Their traditional tasks like offering lectures, seminars, materials, and infrastructure for education will be supplemented by new tasks, e.g. the management of personalized educational programmes.

Unfortunately, the realisation of this scenario is still hindered by the lack of comfortable and easily available technologies supporting students, docents and education managers. The most important technology in this context is the support for search and comparison of study programmes and modules. Such service should help current and prospective students to search for suitable study programmes and single modules as well as to compose personal curricula by expressing their multi-faceted requirements. Managers and docents should be supported when assessing awards of students gained during previous studies, e.g. during a college year abroad. Educational institutions and faculties may use the service for search of potential cooperation partners. The development of a web portal offering a graphical user interface as well as a collection of web services for search and comparison of study programmes and modules, based on Semantic Web technologies, is the objective of the SWAPS project.

To achieve high search effectiveness we set two focal subtasks: 1. Reduction of the user’s efforts needed for multi-faceted search of study programmes and modules through semantic search strategies; 2. Offering users a large range of highly detailed and most recently updated programme and module descriptions by annotating descriptions already available online.

The SWAPS approach exploits the semantics of available programme/module descriptions. Related documents are annotated and indexed with reference to an ontology describing the system of terms and
relations used in the domain of interest. The same ontology is used in the queries set of user’s search operations.

2. Automated annotation of study programme and module descriptions

To make the search results useful and attractive for the users, detailed descriptions of a large range of study programmes and modules must be available. The experience of previous projects [4] shows, however, that it is not realistic to expect that docents and educational managers create suitable dedicated descriptions for the purpose of search systems and that this fact actually builds a strong barrier on the way to an efficient search system. This is the reason why we decided to reuse already existing descriptions of study programmes published on the web.

These documents (such as [2]) describe various attributes of study modules including their allocation within study programmes and their relations to other modules.

The descriptions are usually available in HTML or PDF formats. They are made to be read by humans and are not machine readable so that the information enclosed cannot be extracted and stored in a structured form, e.g. the form of ontologies, automatically. However, information extraction can be carried out by document annotation, a process of marking up particular text chunks representing formally described concepts like person, language, country, time etc.

By now, a fully automatic annotation has not been developed. The state of the art is the semi-automatic annotation: After an annotation is made by machines it is overviewed and adjusted by humans. Yet the existing approaches show a rather high automation level if they are applied to documents using strongly restricted vocabulary and grammar, e.g. if the structure or layout is annotated as described in [3].

Well-defined structure and restricted grammar of programme and module descriptions are essential prerequisites to make descriptions a perfect source for highly automated semantic annotation and indexing. Among 100 evaluated descriptions of study programmes and modules 90 were published in a tabular form, had not more than 40 fields, and a common vocabulary consisting of less than 700 terms.

In the current version of the SWAPS system, the annotation of programme and module descriptions is organized as follows: 1. A docent or a module manager registers a module by submitting either the URL of the online available description or the document itself by means of a special web interface. 2. The automatic part of the document annotation is processed by the system. As the result of annotation, the terms and concepts known by the system are identified. 3. The annotation result is shown to the submitter. (S)he is enabled to change the annotation manually or to change the source document and to restart the annotation process. 4. The extracted data are stored in the SWAPS repository and used for indexing as well as for extension of the domain ontology.

3. Perspectives

By now, the kernel components of the SWAPS system (the search engine, the interface for the submission of descriptions and the annotation tool) have been basically developed. For system deployment on a large scale we need a critical mass of module and programme descriptions. The large number of related descriptions to be found on the web, however, cannot be assumed to be reliable and up-to-date. To guarantee the reliability of information, we set up a procedure for submission of these descriptions by docents or by faculty staff.

To this concern we are looking for cooperation with faculties and departments of universities. Starting within a local network of German universities, we wish to extend partnership and cooperation to international partners in the near future.

4. References