



Project no. 027391

Project acronym: LT4eL

Project title: Language Technology for eLearning

Instrument Specific Targeted Research Project

Thematic Priority Information Society Technology

D6.2 Exploitation plan (update) and report

Due date of deliverable: 30-11-2006

Actual submission date: 21-12-2006

Start date of project: 1-12-2005

Duration: 30 Months

Organisation name of lead contractor for this deliverable: Utrecht University (UU)

Revision [1]

Project co-funded by the European Commission within the Sixth Framework Programme (2002-2006)		
Dissemination Level		
PU	Public	x
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium (including the Commission Services)	

D6.2 - Exploitation plan and report

Contents

- 1 Executive Summary
- 2 Introduction
 - 2.1 Exploitable results
 - 2.2 Potential users
 - 2.3 Main exploitation activities
 - 2.3.1 Dissemination
 - 2.3.2 Marketing
 - 2.4 IPR issues
 - 2.5 Division of tasks
- 3 Exploitation plan
 - 3.1 Know-how
 - 3.2 Corpus of learning objects
 - 3.2.1 Description
 - 3.2.2 Use and exploitation
 - 3.3 Tools for metadata generation: key word extraction, glossary candidate detector
 - 3.3.1 Description
 - 3.3.2 Use and exploitation
 - 3.4 Domain specific ontology and language specific vocabularies
 - 3.4.1 Description
 - 3.4.2 Use and exploitation
 - 3.5 ILIAS integrated functionalities
 - 3.5.1 Description
 - 3.5.2 Exploitation and Use
 - 3.6 Validation methodology
 - 3.6.1 Description and exploitation
- 4 Overview table

Executive Summary

This document describes the LT4eL project exploitation plan. The purpose of this document is to describe how the results of the projects will be used or exploited by the LT4eL partners and by other potential users during the remaining months of the project as well as after the project. The document is organized as follows:

- The first section identifies the exploitable results, the potential users, and the main exploitation activities.
- The second section contains the actual plan describing, **per exploitable result**, how it can be exploited or use in further research.

Introduction

This introductory section identifies and briefly describes the exploitable results, the potential users, the main exploitation activities and IPR protection measures.

Exploitable results

- Know-how in Language Technology (LT), Natural Language Processing (NLP), Semantic Web and eLearning technologies;
- Corpus of learning material of at least 1000 pages for each language represented in the consortium (Work Package 1);
- Tools (methodology) for metadata generation (key word extractor, glossary candidate detector) (WP2);
- Developed domain specific ontology with at least 1000 concepts, and language specific vocabularies (WP3);
- ILIAS integrated functionalities (keyword extractor, glossary candidate detector and multilingual retrieval via ontologies) (WP4);
- Validation methodology (WP5).

Potential users

- LT4eL partners who will use the collected knowledge and resources for their research activities and the developed functionality integrated in the ILIAS system for their didactic activities.
- Scientific community (LT, NLP, Semantic Web and eLearning), including other ICT projects, who will use the collected resources, the developed concepts and methodology for their research activities.
- LMS developers who will use the developed concepts and methodology to integrate new functionalities within their systems.
- LMS end-users who will use the new functionalities integrated in the ILIAS or other systems.
- Local authorities (e.g. ministries of education), and higher education institutions (e.g. universities) who will become aware of the potential of eLearning.

Main exploitation activities

Dissemination

Dissemination of the results will be ensured through the multilingual Web portal, participation at scientific events, notably LT, NLP, Semantic Web and eLearning conferences and workshops, as well as by means of reports, mailing lists and the project newsletter. In order to promote the use of the LT4eL functionalities, and in general of open source Learning Management Systems (LMS), we will establish contacts with local educational authorities (ministries of educations, universities and schools) in particular in the new member states, for example via the organization of special meetings or awareness events on a national level. The technical results of the project will be made available in well-documented releases. The new functionalities as well as the enhanced LMS will be freely available under the GNU General Public Licence (GPL) and distributed also through the ILIAS channels, as well as through the "Language GRID" initiative in which the University of Tübingen participates. For a more detailed description of our planned dissemination activities see deliverable D6.1.

Marketing

The functionalities developed within the project will be integrated in the ILIAS Learning Management System (LMS). We will however market our approach with other open source or commercial systems as well, such as Moodle, OLAT, ML3, ATutor, Sakai, Claroline, Blackboard, WebCT. To this aim a research about content formats and metadata support used in other LMSs has been already carried out and published in the project internal webpage. To create interest of potential users and prospects, as well as direct contact with other LMS developers, the consortium plans to participate to various eLearning or other LMS-related exhibitions and

demonstrate the developed tools and methodology. The following is a list of exhibits where the LT4eL tools and developed methodology could be demonstrated:

- EURODIDAC (<http://www.worlddidac.org/eurodidac/>)
- Online Educa (<http://www.online-educa.com/> (Berlin)
<http://www.online-educa-madrid.com/> (Madrid))
- Learntec (<http://www.learntec.de/>)
- ECEL EUROPEAN CONFERENCE ON eLEARNING
(<http://www.academic-conferences.org/ecel/ecel2007/ecel07-home.htm>)
- International Conference on Information Communication Technologies in Education.
(ICICTE) (<http://www.icicte.com/isite/home/>)
- European Conference on Educational Research (<http://www.eera.ac.uk>)
- European Conference on Information Systems (ECIS) (<http://www.ecis2007.ch/>)
- German e-Science Conference (<http://www.ges2007.de/>)

OU, ZHW and the ILIAS team will act as link with the eLearning community. For example, the ILIAS team is part of a German network of eLearning developers called Campussource (<http://www.campussource.de/org/>), and the eLearning Centre of the ZHW has direct contact with Moodle, OLAT and ML3. Both channels will be used as starting point to establish collaborations.

UU has further established contacts with Giunti Labs who is monitoring the project in order to assess potentiality for commercialization of the functionalities developed.

Finally, to further favor exploitation, the developed functionalities will also be offered independently as web services so that they can be used on the internet by other applications as well. And the project decided to make the source code available under an open source licence and to host it on the SourceForge.net portal for open source projects at <https://sourceforge.net/projects/lt4el/>. This makes the results immediately available to the general public and gives everyone the opportunity to join and collaborate with the project.

IPR issues

All the IPRs (Intellectual Property Rights) concerning publications and other know-how resulting from the project will be handled in agreement with all partners. A budget is foreseen in case it is necessary to cover IPR for the learning objects collected. The IPR status of these documents is still under investigation (see activity report). It is the task of the WP-managers that all issues related to the intellectual property will be respected. Basically our project will deal only with freely available resource and tools. However, sources and authors of all resources, tools and platforms will be mentioned in the web inventories and in the reports. The project deliverables will be publicly available. Reports of the consortium management boards will be available only for the project members.

Division of tasks

Exploitation activities are part of WP6 with Utrecht University as WP leader. All partners however are committed to the widespread dissemination and active exploitation of the results of the project.

Exploitation plan

The exploitation plan consists of a short text per exploitable result including a description of the result and a plan for its use or exploitation. Since most of the exploitable material is still in preparation, this plan is necessarily incomplete, and will be integrated as the project evolves.

Know-how

Through the LT4eL project, each partner acquires knowledge and skills in state-of-the-art technologies in LT, NLP, Semantic Web and eLearning. It is only natural for all partners to exploit this know-how for their future research and for the acquisition and execution of future projects.

Corpus of learning objects

Description

This corpus consists of at least 200000 words (1000 pages) of learning objects for each language represented in the consortium. The broad domain of this collection is the use of computers in education. More specific subdomains include teaching academic skills, creating webpages, basic computer skills. This material has been linguistically annotated at least up to the level of part-of-speech tagging and morphological analysis, and has been marked up with (i) key words, and (ii) definitory contexts. The complete annotated corpus is delivered at month 12 (December 2006).

Use and exploitation

The relevance of this corpus lies essentially in its use for the development of the new functionalities within the project. However once all IPR issues will be resolved the complete corpus will be made publicly available in the ILIAS library, but also on the project portal. As potential users we identify ILIAS end-users (teachers and students) interested in the content of the documents, but also the whole scientific community, notably other ICT projects like iCamp (<http://www.icamp-project.org/>). The latter can use the multilingual corpus - which, if necessary, could be easily annotated with other levels of annotation - as a data base for further experimental research, for example, in language technology and natural language processing. Knowledge about the availability of this corpus will be disseminated via the usual channels (project webportal, project eNewsletter, presentations at scientific events).

Tools for metadata generation: key word extraction, glossary candidate detector

Description

The aim of the project is to improve the retrieval and accessibility of content through the identification of the learning material by means of descriptive metadata. To this end, we employ available language technology tools and resources to develop two functionalities that facilitate the semi-automatic generation of metadata. The two functionalities are:

- A keyword detector and extractor, i.e. a tool that supports authors and content managers in selecting, in the chosen learning objects, the keywords that best represent the topic(s) of these learning objects. The tool analyses a set of annotated documents and returns the best keyword candidates for each learning object. The user of this functionality decides on the inclusion of these candidates into the metadata.
- A glossary candidate detector, i.e. a tool that supports authors and content managers in creating glossaries by identifying definitory contexts in a text that contains the term to be defined and its definition. Again the input of this tool are annotated texts (and background resources). The output of this tool is a set of candidates for definitions. Again the user decides on the inclusion of these candidates into the metadata.

We expect that the addition of these functionalities will dramatically increase the retrievability of the learning objects in terms of their content. The tools will be fully documented. The documentation is essential not only for a proper integration within the ILIAS system, but also for the integration within other LSM that will be interested in adopting them. The first version of the tools is delivered at month 12 (December 2006). The final version will be delivered at

month 30 (May 2008).

Use and exploitation

The developed functionalities will, in the course of the project, be integrated into the Learning Management System ILIAS. To favor further exploitation, they will also be offered independently as web services so that they can be used on the internet by other applications as well. Full integration into other LMSs is also possible through the open interfaces of the tools and their full documentation. To foster implementation into other systems, a document describing the integration process into ILIAS (a step-by-step integration manual) will be provided targeted at developers of other LMSs, and a workshop on the integration process will be organized in 2008. The consortium further plans to participate to eLearning or other LMS-related exhibitions to demonstrate the developed tools and methodology as well as to create direct contact with other LMS developers (see section *Marketing* above for a description of other promotional activities). Beyond direct exploitation in the eLearning industry, we foresee for these results an impact on the scientific community. These functionalities use state-of-the-art LT and NLP technologies and therefore can provide useful feedbacks to these communities and be the starting point for utilization of these technologies into other domains. To further promote the use of NLP tools in eLearning and other fields, a workshop on 'Natural Language Processing for Metadata Extraction' has been organized at the The Twelfth International Conference on Artificial Intelligence: Methodology, Systems, Applications (Aimsa) (Bulgaria, 13-15 Sept. 2006) (<http://www.bultreebank.org/NLP4ME2006/>).

Domain specific ontology and language specific vocabularies

Description

The domain specific ontology (i.e. use of computers for non specialists) is developed in a language independent way and comprises round 1000 concepts. An English vocabulary is mapped to the concepts and the relations within the ontology. In a later stage, starting from January 2007, language specific vocabularies will be developed and will be linked to the ontology for all the languages of the consortium. The domain specific ontology is delivered at month 12 (December 2006), and the language specific vocabularies at month 18 (May 2007). Ontologies will be used to structure, query and navigate through the learning objects that are part of a LMS. The ontology can play two major roles:

- Classification of learning objects. Each learning object is connected to a set of concepts in the ontology. This classification allows ontological search, i.e. search based on concepts and their interrelations within the ontology.
- Multilingual search for learning objects. In this case the ontology plays the role of Interlingua between the different languages. Thus the user might specify the query in one language and get learning objects in other language(s).

The innovative aspects of this part of the project consist firstly in the application of semantic web technologies (ontologies) to facilitate learning processes and, secondly, in their use (linked to language specific vocabularies) to address problems of multilingual nature, in particular multilingual search.

Use and exploitation

From the perspective of our project, the relevance of these semantic web technologies lies essentially in their use within an LMS to improve the organization and the retrieval of the learning objects across languages. Potential users in this respect will be LMS (notably ILIAS) end-users, but mostly LMS developers, open source as well as commercial. Documentation of the ontology development and report for its integration within an LMS will be provided to facilitate its utilization in any eLearning platform. However, again, we expect also a scientific impact of these results. The use of ontologies for multilingual retrieval is in its infancy. We

expect to contribute to the development of these techniques (to this aim we are in contact with the APOSDLE (<http://www.aposdle.tugraz.at/>) project to use common resources) and to provide important feedbacks to the semantic web community on the potential of their technology in this field. To further promote the use of semantic knowledge in eLearning we plan to organize a workshop on this topic at the EUROLAN 2007 which will be held in Romania in the summer 2007, and we have been and will be present at various Semantic web and eLearning events. The developed ontologies and in particular the methodology to link these to language specific vocabularies will be further disseminated via the usual channels to be used by the scientific and semantic web community for further research.

ILIAS integrated functionalities

Description

The tools described above (key word extraction, glossary candidate detector and ontologies for multilingual retrieval) will be implemented within the ILIAS system. ILIAS is a web-based learning management system and allows users to create, edit and publish learning and teaching material in an integrated system with a normal web browser. Tools for cooperative working and communication are included as well. ILIAS is available as open source software under the GNU General Public License. The software development worldwide is coordinated by the team at the ETHZ. The current release 3.5.0 of ILIAS already offers content authors the possibility of annotating learning objects with metadata based on the LOM standard. Also a metadata based search is available. However, the meta data have to be provided manually by the author. This makes the annotation process very time consuming and thus only a few authors provide useful metadata. Tools for semi-automatic metadata generation could help solve this problem. Furthermore ILIAS does not provide semantic web based functionalities but it already offers the possibility of reusing learning objects like media objects or glossary items in the creation process of learning material. Ontology based retrieval of learning objects will considerably improve the task of reusing learning objects since ontologies will allow for intelligent searching and navigation in huge amounts of data. Metadata annotation and ontology driven search and navigation will allow for individual content assemblance for learners. Learners will be able to build individual learning paths by entering key terms of concepts they need to learn.

Exploitation and Use

The main users of these integrated functionalities are ILIAS end-users (teachers and students). See the appendices of deliverables D2.2, D2.3 and D3.1 for a detailed description of use cases related to the tasks performed within ILIAS that use the tools provided by the LT4eL project. To promote the use of the newly developed functionalities, the ILIAS English user manual will be extended with detailed descriptions of these functionalities. The user documentation will be distributed online via the ILIAS website as well as via the project website.

Validation methodology

Description and exploitation

A suitable validation methodology will be developed which will be applied to the validation of the new functionalities as well as to their integrated set into ILIAS. eLearning applications are very much an emerging field, and there are no standard, general methodologies that can be used to validate effectiveness of the learning process in our specific context. We expect the methodology developed within the LT4eL project to be at least a first step towards this missing standard. To promote dissemination, and thus exploitation, a report on the developed validation methodology will be produced at Month 18 (May 2007) and made public on the project portal.

Overview table

Exploitable Material	Sector of Application	Timetable for use	Patent or other IPR protection	Owners and other partner involved
Corpus of LOs	Research & eLearning industry	2007	IPR free	all partners
key word extraction & glossary candidate detector	Research & eLearning industry	first version: 2006; final version: 2008	IPR/patent free	all partners
ontology & language specific vocabularies	Research & eLearning industry	2007	IPR/patent free	all partners
ILIAS integrated functionalities	eLearning	first version: 2007; final version: 2008	IPR/patent free	all partners
Validation methodology	Research + eLearning industry	2008	IPR/patent free	all partners