Development-oriented Open Source eLearning Tool Evaluation: the Edukalibre Approach

Luca Botturi
University of Lugano
Switzerland
luca.botturi@lu.unisi.ch

Chris Tebb
University of Leeds
UK
chrispy@comp.leeds.ac.uk

Vania Dimitrova
University of Leeds
UK
vania@comp.leeds.ac.uk

Drew Withworth
University of Leeds
UK
drew@comp.leeds.ac.uk

Julika Matravers
University of Leeds
UK
julika@comp.leeds.ac.uk

Jutta Geldermann
University of Karlsruhe
Germany
jutta.geldermann@wiwi.uni-karlsruhe.de

Isabelle Hubert
University of Karlsruhe
Germany
isabelle.hubert@wiwi.uni-karlsruhe.de

I. INTRODUCTION

Numerous eLearning tools have been developed to date. They promise to enrich the learning experience by offering flexible access to, and sharing of, resources and communication between students and tutors. The increasing amount of Open Source eLearning solutions, in particular, has recently provided the educational sector with access to flexible alternatives to the traditionally expensive commercial products which many institutions cannot afford.

With the rapid invasion of these technologies into educational institutions, the evaluation of these systems has become increasingly important. Their evaluation is required in order to ensure further research and development through the provision of feedback for further improvement of these systems. The Open Development Model, in particular, relies on the improvement of software products through its modification by a wide range of developers including both teachers and students. The provision of constructive feedback is crucial to the success of such a novel development mentality that is highly iterative with notably frequent software releases and a geographically distributed development team. However, the development of a suitable framework for the comprehensive evaluation of such eLearning applications remains a challenge.

Typically eLearning tools have been evaluated on the basis of particular aspects of a system such as their costs, their technical specification, the features they provide and their usability [1]. The importance of usability evaluation, in particular, is frequently stressed in the literature. Magoulas et al [2], for example, propose the use of heuristic evaluation within a layered evaluation approach for usability testing within Adaptive Learning Environments.

Research also seems to have given priority to the evaluation of system usability against the background that the evaluation of the effectiveness of a learning environment remains difficult. Furthermore, it can be argued that good usability is an essential feature of a successful eLearning system. The level of a system's usability may not only influence the user acceptance of a system but may also greatly influence the learning process [3]. Whilst it needs to be recognised that good system usability is crucial for the user acceptance of an eLearning system, it needs to be acknowledged that usability is a necessary, but not a sufficient, condition for an effective learning environment [4].

It is the objective of this paper to propose an evaluation approach for eLearning applications that not only recognizes the need for a high level of a system's usability, but at the same time opens up scope for a more holistic evaluation approach that accounts for, and supports, the rapid evolution of Open Source Software and it is suited to the peculiar development process of libre applications. Such an evaluation approach addresses both the behavioural aspects and the system's architecture, thereby aiming to support the rapid development and refinement of eLearning systems by producing design-oriented feedback for developers.

II. THE EDUKALIBRE EVALUATION APPROACH

A. About Edukalibre

Examining the connection between libre software development and creation of open content for education is the main objective of the EU funded Edukalibre project [5] [6] that involves developers and educationalists from several European countries working together to study both technological and pedagogical aspects of the successful deployment of the libre idea in university teaching. Edukalibre is a project to explore the open source development methodology (geographicaly distributed authors, version controlled authoring, etc..) and its application to the collaborative creation of academic content.

The core of the system is the repository manager, written in python, which uses the subversion version control system. The system accepts documents in openoffice, latek or docbook formats, and will automatically version control the document and provide output versions in multiple formats (currently xhtml, docbook XML, text and PDF). This repository manager has been designed to be independant of the user facing applications, allowing us to use openoffice as our interface (working directly with the webdav repository) or web based collaboration software. We have created two web based systems, one simple document management tool called COLLAB and COnDOR, a groupware tool which has been created as a module of the open source VLE Moodle (see http://moodle.org). Moreover, a visual user-tracking tool was developed, again for the Moodle environment, called GISMO [7].

The main idea behind Edukalibre is to provide tools that enable the deployment of the libre development model in educational settings. It is expected that the tools

developed will facilitate the collaborative construction of educational resources by both teachers and students. The specifics of the open development model brought some challenges that had to be dealt with during the evaluation.

B. Evaluation Goals for Edukalibre

In the first place, the whole design and development cycle was iterative, including frequent release of software that was sometimes not fully tested. Moreover, the development, as well as the evaluation, involved several geographically distributed teams. The advantage of this is the quick deployment in various settings and the wide dissemination. The downside is that the effort has to be distributed and the work properly synchronised. As a result, several small scale evaluative studies appear to be more appropriate than one large evaluation. Furthermore, the development was done asynchronously at times with several developer teams working independently in several countries (Spain, Leeds, Switzerland, and Portugal). This led to the need to be able to conduct evaluation in a flexible way to enable examining different tools within the appropriate context. These constraints had to be taken into account in conducting the Edukalibre evaluation, as discussed in the following sections.

Another issue that challenged the definition of an evaluation approach for Edukalibre was the relationship between software application design and development and the evaluation itself. A complete evaluation that looks for comprehensive data, requires real users in real setting and stable tools in their final version. Such an approach produces sound *ex-post* results which might be used to promote the tools or to improve the design in the future. Edukalibre, given its libre nature, had different needs: on the one hand, developers needed short term input and guidance relating to their development work in order to improve their tools; on the other, the project team felt the need of a wide evaluation of the impact of the tools in a real higher education setting in order to evaluate the usefulness and effectiveness of their product.

C. Evaluation outline

Based on these goals and taking into account the specific needs of the libre development model, we developed a three-phase evaluation approach.

Phase 1 included several studies involving experts and users from different partner institutions focused on usability. A usability framework was developed following the guidelines of the MiLE method (Milano-Lugano Evaluation method [8]), and was shared amongst all evaluation teams, so that results were then comparable. This method was then combined with cognitive walkthrough involving potential student users. Phase 1 was conducted between October and November 2004.

Evaluation can be addressed properly only if inspection is combined with real user trials. Phase 2 focused on student feedback, in order to verify the results of phase 1 and to investigate the reasons for the problems detected during the usability inspection. In line with the Edukalibre objectives, the initial user trials consider the creation of open educational content by both user groups - teachers and students. Phase 2 was conducted in December 2004.

Phase 1 and 2 should then be complemented with a broader holistic evaluation of the impact of the system in a real educational setting – and this exactly the goal of Phase 3, scheduled for Spring 2005.

Figure 1 represents the temporal disposition of the three evaluation phases and their relationships with each other and with the development process.

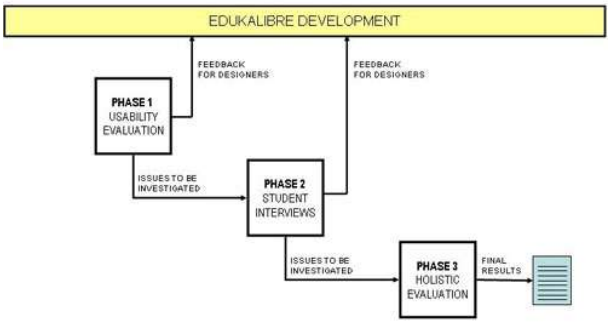

Fig. 1 - Evaluation approach outline

III. THE EVALUATION PROCESS

A. Phase 1: Rapid Usability Inspection

Usability evaluation is a narrow-scope assessment of how satisfactorily users can achieve specific goals with a web application. With the Edukalibre framwork, usability evalaution was selected as (a) it can be conducted in a short time, so to keep the pace with the needs of the developers; (b) it produces design-oriented feedback useful for the refinement of the application; and (c) it can be conducted in a uniform way by different teams in a libre fashion, if a specific methodology is followed.

In order to meet the conditions expressed under (c), the MiLE method was selected. MiLE is an experience-based evaluation framework for web applications that strikes a healthy balance between heuristic evaluation and task-driven techniques. MiLE can be used flexibly at different levels of granularity, according to the resources available to the evaluators. It offers reusable tools and procedures to carry out both expert reviews (called *inspections*) and user testing within budget and time constraints.

Given the goals for the Edukalibre project expressed above, MiLE was slightly adapted. The usability framework (i.e., application requirements, scenarios, goals and tasks) was developed from the application requirements expressed by the developers' documentation (which were in turn gathered from user requirements). Scenarios were developed both for teachers and students. This information was distributed to all project partners, along with indications about how to conduct the inspection. Each partner was in charge of assessing the usability of its module and of one other. The results were collected with a standard usability matrix, to be sent to the Lugano team, which was in charge of collating the data and producing the final guidelines. The Leeds team set up Moodle courses that reflected the scenarios and were used for inspection. In the end, GISMO was evaluated by 3 inspectors on a single teacher scenario, COnDOR by 4 inspectors on 2 teacher scenarios and by 3 other inspectors on 2 student scenarios.

In order to provide the flavour of the results obtained with MiLE in Edukalibre, and their direct impact on design, two sample results are reported. Inspecting GISMO, an issue emerged concerning orientation and

navigation for Task 02: See who logged in during last week. Actually, in order to see logins, the teacher should select *draw chart* and then *login overview*, which is not very straightforward. The identified solution is to organise menu semantically (i.e., following the teacher's perspective), and not operationally (i.e., following the programmer's perspective).

For COnDOR, an issue was identified concerning Task 08: Identify the least and most productive contributors of Group 1 (a group in the scenario), concerning the content dimension. Actually, the COnDOR interface is document-oriented, and in order to track the activity of single students, the teacher had to browse and count messages/posts manually. This was indeed a common task among teachers, which was not supported by the tool. It was therefore included in the plan for next development to provide summary information about the activities of each group member (including the number of posts and file contributions, the date of the last contribution, etc.).

Usability evaluation provided useful indications for further development – yet it is by nature not complete, as it only focuses on a specific feature of the system, and it does not consider its educational impact and its effectiveness. This is the reason why Edukalibre moved forward to the next two phases.

B. Phase 2: Student Feedback

In order to verify the results of the rapid usability inspection (Phase 1), student feedback was gathered in phase 2. At the University of Karlsruhe, a seminar was organized with 14 German students in the advanced study period of Industrial Engineering. During four months, the students elaborated an environmental business game, which comprised computer supported case studies for location planning and technique assessment. Four groups of three students worked on a specified topic, representing fictitious companies faced with strategic questions and the subsequent decisions. In addition, two individual topics were related to business games and open source software [9].

During the preparation of the business game, the students were invited to use the Edukalibre platform that was specifically set up for them. The platform comprised of the general moodle platform (with news forum, calendar, documents etc.) plus the developments of Edukalibre up to that date: one groupware tool for each group as well as a direct link to the conversion platform.

For producing development-oriented feedback, a questionnaire was used with 15 questions, developed according to the issues identified in Phase 1. The results confirmed some results obtained in Phase 1, and showed that the students were not totally convinced of the benefits of any eLearning platform mainly because they can meet regularly and attend “traditional lectures”. They did not use the platform regularly, which explains most of the lukewarm answers found in this first questionnaire. More documentation such as a tutorial for the Edukalibre-platform would be beneficial. Yet, the idea is growing and the acceptance and extensive use of such a platform by the students could arise soon.

C. Phase 3: Holistic Evaluation

The next phase of Edukalibre's evaluation, is scheduled for Spring 2005, and it complements Phases 1 and 2. The main idea is to investigate the developers and find out if and how their assumptions match with those of the users of their applications – students and teachers. This will alllow a comprehensive final evaluation of the impact of Edukalibre tools in the educational practice.

Why is it necessary to investigate the developers? Although powerful for delivering results during development, we believe that standard HCI recommendations for usability testing are inherently limited in terms of what they can reveal. HCI can ask how well "objective" usability standards are being maintained, how accurately the needs of users are being assessed and how effectively the insights gained are translated into system functionality. However, usability problems are, at heart, conflicts between the assumptions of the developers and the needs of the users. These may emerge because the needs of users have not been adequately researched: here, traditional HCI (as summarised above) can help. However, they may also emerge because of inherent differences in how different groups perceive a system, and this can apply both to the specifics of a system (functionality), and generalities (such as why the system was created in the first place). Here it is not a case of lacking the right answers; evaluators may not even be asking the right questions.

In the first place, many usability evaluations do not take place in the full learning context [10]. Secondly, task-driven evaluation is limited to assumptions made by the designer so will struggle to test unexpected and creative use of the system. Nor do these methods consider technologies as being dynamic, existing in a constant state of evolution [11].

This is not meant to tar Edukalibre with the brush of commercialism: it is merely an example of the potential ways in which eLearning technologies, being complex and significant to a wide variety of stakeholders [12], can become an arena which different interests and perspectives compete to define. The "social shaping of technology" thesis [13] suggests this; recent work on the politics of cognition [14] has also suggested that many of the assumptions which become encoded into technologies are unvoiced or unconscious, shaping not only the answers to certain questions but defining what questions can be asked in the first place.

The ideal is first that assumptions are revealed by self-reflection; second, that they can if necessary be challenged and changed. This has been termed "double-loop learning" [15]. Ideally, this takes place in a continuous cycle of evaluation then subsequently, implementation. The more user needs can be included and underlying assumptions revealed, the more participatory the construction of the environment will become. This is a driving force behind the project in the first place (thanks to its basis in open source software and collaborative content) so the method is appropriate to the system.

IV. CONCLUSIONS

Open Source eLearning applications have become increasingly common in today's educational environment. However, the lack of suitable evaluation approaches - that extend beyond the assessment of a system's usability - to provide developers of eLearning applications with constructive feedback remains an issue. This issue now

extends further into a world where the Open Development Model has become a dominant development mentality.

It was the objective of this paper to propose an evaluation approach for eLearning systems that follow the Open Development Model. It was argued that such an evaluation approach would have to be holistic in the sense that it offers feedback beyond issues of usability. Furthermore, it should be applicable within a development environment that was highly iterative with frequent releases, and that potentially involved a high number of geographically dispersed developers of various backgrounds (including students and teachers).

It was not the intention of this paper to provide a recipe for the complete evaluation of eLearning systems. However, this paper recognizes that software development projects normally have an ambitious time frame, and results of ex-post evaluation can hardly be taken into account in full. The evaluation approach proposed here supports a holistic evaluation of Open Source Software that accounts for these factors. This paper arose from the recognition that a lot more work is required to address the issue of suitable and complete evaluation for Libre Software Development. The evaluation approach proposed here offers a significant step towards addressing these emerging needs.

V. REFERENCES

[1] Britain, S., & Liber, O. (2000). *A Framework for the Pedagogical Evaluation of Virtual Learning Environments*. JTAP Report No. 041.

[2] Magoulas, G. D., Chen, S. Y., & Papanikolaou, K. A. (2003). Integrating Layered and Heuristic Evaluation for Adaptive Learning Environments. In S. Weibelzahl and A. Paramythis (eds.), *Proceedings of the Second Workshop on Empirical Evaluation of Adaptive Systems*, held at the 9th International Conference on User Modeling UM2003, Pittsburgh, 5-14.

[3] Strother, J. (2002). An Assessment of the Effectiveness of Elearning in Corporate Training Programs. *International Review of Research in Open and Distance Learning*, 3(1).

[4] Pulichino, J. (2004). *Usability and eLearning*. eLearningGuild Report.

[5] Gonzalez-Barahona, J. M., Dimitrova, V., Chaparro, D., & Mazza, R. (2005). Producing educational resources in the libre way: The Edukalibre project. Forthcoming.

[6] Edukalibre (2005). Edukalibre's Website. Retrived on February 23rd, 2005, from www.edukalibre.org.

[7] Mazza, R., & Milani, C. (2004). GISMO: a Graphical Interactive Student Monitoring System for Moodle. Paper presented at the *UK and International Moodle User Conference - MoodleMoot 04*, Oxford, UK.

[8] Triacca, L., Bolchini, D., Botturi, L., & Inversini, A. (2004). MiLE: Systematic Usability Evaluation for Elearning Web Applications. *EDMEDIA 2004, Lugano, Switzerland,* 4398-4405.

[9] Geldermann, J., Treitz, M., & Rentz, O.(2004). Teaching Integrated Technique Assessment based on Open-Source Software. *EUNITE Conference: European Symposium on Intelligent Technologies, Hybrid Systems and their implementation on Smart Adaptive Systems, Aachen, Germany, 10.-12.06.2004,* PP 503-508.

[10] Lim, K. H., Benbasat, I. & Todd, P. A. (1996). An experimental investigation of the interactive effects of interface style, instructions and task familiarity on user performance. *ACM Trans. Computer-Human Interaction*, 3(1), 1-37.

[11] Star, S. L. (1999). The Ethnography of Infrastructure. *American Behavioral Scientist*, 43(3), 377-391.

[12] Benson, A. (2002). Using Online Learning to Meet Workforce Demand: A Case Study of Stakeholder Influence. *Quarterly Review of Distance Education*, 3 (4), 443-452.

[13] Mackenzie, D., & Wajcman, J. (Eds.) (1985). *The Social Shaping of Technology.* Buckingham, UK: Open University Press.

[14] Blaug, R. (2005) *Power and Cognition*, forthcoming.

[15] Argyris, C. (1999). *On Organizational Learning.* Oxford, UK: Blackwell.