Visualizing Learner Models through data aggregation: a test case

Luca Mazzola, Riccardo Mazza

Abstract: The needs of tools able to effectively support the learning process is a well known open issue in the field of Technology Enhanced Learning. Two different directions for improving this situation could be explored, depending on the pedagogic model: directly support the self-reflection in a more self-regulated learning experience or fostering the tutoring processes for the supervision of learners' activities. Stressing the data collected in TEL platforms by the processes of adaptation and personalization, typical of Web 2.0 and its interaction model, we developed an infrastructure for creating adaptive visualization of learners' profile in the context of a distributed and heterogeneous environment. In this article we present an initial evaluation of the software: after introducing the structure of the tool, called GVIS, we present a first test case with very simple representations in a real course and the users' feedback about their experience. A set of more advanced visualizations are under development and we expect to perform further evaluations using a questionnaire on the new mockups. The data already collected showed a good user experiences, although some problematic aspects still exist and highlights issues to still be addressed.

Keywords: Technology Enhanced Learning, Open Learner Models, Adaptivity, Information Visualizations, Data Mashups, Life Long Learning

Introduction

The needs of tools able of effectively supporting the learning process is a well known open issue in the field of Technology Enhanced Learning. This is particularly important due to the limitations in communication channel that are related to the mediation by technologies and the removal of physical presence. Two different directions for improving this situation could be explored, depending on the pedagogic model: directly support the self-reflection in a more self-regulated learning experience or fostering the tutoring processes for the supervision of learners' activities. Both of them could be effectively supported through the creation of graphical representation of relevant data, aggregated at different level and with different view, based on the roles and the profile itself.

Furthermore, another component play a major role: the paradigm of Web 2.0 has changed not only the way we explore and search in the Internet network, but also how users expect to interact with online resources. Posting comments about a news in a blog, refining an article in a collaborative wiki, or aggregating information from heterogeneous sources are features that also affect the field of eLearning, thus determining the space for a wider definition of Technology Enhanced Learning (TEL). Instead of trying hard to accurately reproduce the face-to-face experience in an online environment, recently a new approach has emerged: its main characteristic could be seen in the effort of stressing what the electronic medium offers and the technologies implemented on it. As a consequence, the introduction in the learning process of different tools, platforms, widgets and devices creates a personal space for a potential fruitful usage of the rich and diffused amount of resources available for the learner experience. The availability of many contents not specifically designed for being part of a structured flow, enriches the learning experience and demonstrates the potential impact of Informal Learning. In such open approach, the concepts of *personalization* and *adaptation* assume a central role. On the personalization side, the possibility to choose among different options – such as the type of media used (text, audio or video) or the approach adopted in the subject presentation (inductive or deductive) – allows learners to have a well-suited experience. On the adaptation side, the students are provided with the content that is appropriate to their profile. The integration between many and dynamically added sources, heterogeneous tools, devices for different situations, supported with the two processes of personalization and adaptation, creates a working space known as Personal Learning Environment. In this context researches have demonstrated that it is useful to consider the richness of the experiences adopting a holistic approach (McCalla 2004), such as it is normally done with complex system, like a natural ecosystem in ecological analysis.

For these adaptive features, one of the most important component is the *student model*, which is in charge of keeping track of the learner's knowledge and skills acquired during the learning process. As already discovered (Bull & Kay 2008), in order to increase the level of engagement of learners, stimulate the perception of their current status, and to encourage reflection as learning (Bull 1997), the student model can be opened to the inspection of learners and instructors. Student models are usually made available as visual representations, because it simplifies the data interpretation. Since the user information is not only stored into a specific student model system, but is often distributed in a number of platforms used for different purposes, data must be aggregated from different tools (e.g. student models, intranet usage data, LMS data) and provided consistently to the interested users, preferably in visual format (Dror et al. 2008). We designed a specific multi-tier infrastructure for this purposes, in the context of the EU-FP7 GRAPPLE project (Van Der Sluijs & Hover 2009), which aims at making already existing LMSes able to include adaptive contents and at sharing some controlled set of data about users. Other approaches to the problem of creating personalized experience in TEL are "educational mashup" (Esposito et al. 2004) and "ubiquitous and decentralized user model" (Van Der Sluijs & Houben 2006, Heckmann & al. 2005). In these works researchers aims at aggregating user data from different systems, but they tackle to problem from different point of view. From the user point of view, we have to consider that a high quantity of mashed-up data might cause an overload problem (Chen 2009), that becomes problematic when it distracts from the learning activity and makes the learner confused about the data represented. Our solution consists of compact and detailed representations. We provide mash-up data as compact indicators in the main LMS user interface, while the detailed view are represented through set of widgets in the user dashboard. To limit the overload problem, the visual representations can be made adaptive to the role, to the context, and to the activities performed by the learner. The adaptation helps in creating more comprehensible and easier indicators. For example, the adaptive dashboards adopted in the field of Business Intelligence (Schutz 2009) are used to representing the most useful and relevant subset of the all information available for the ongoing task.

Next section will describe the infrastructure we implemented; then we will discuss how the adaptable features provided could offer interesting capability in the context of Web 2.0 and of mashuped data; then we present an initial implementation with very simple indicators we run in the context of a real course and the survey we carried out with the involved learner. We conclude the article presenting some reflections on the collected data and what we expect as output from the application of our infrastructure in TEL projects.

The Infrastructure

GVIS - acronym for GRAPPLE Visualization Infrastructure Service - is the infrastructure we developed to extract data from different sources and enable instructional designers to easily create adaptive indicators of the learning state for learners and tutors.

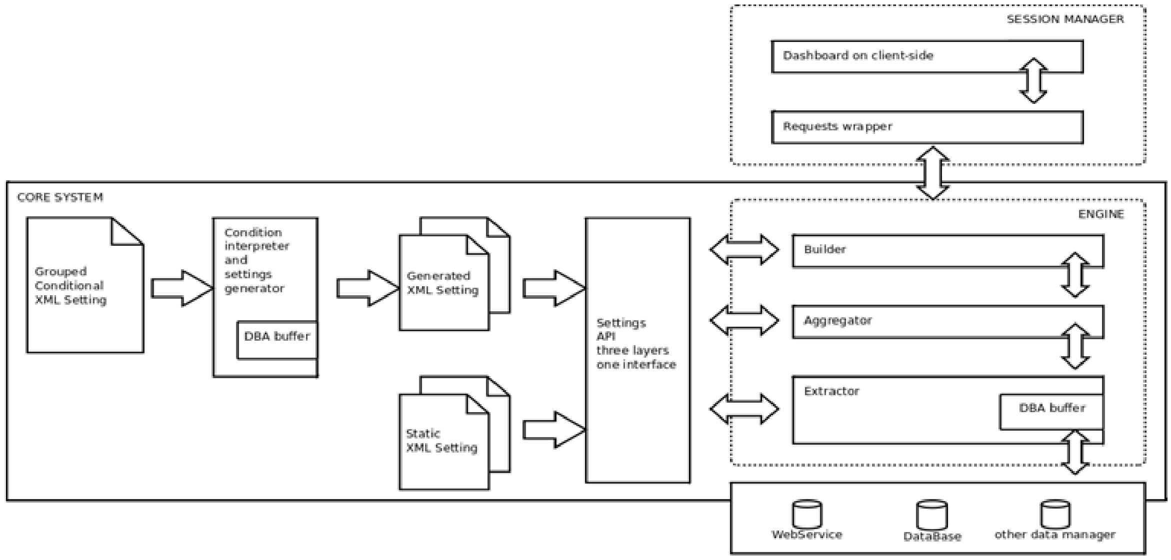

Fig. 1 The GVIS architecture: the adaptation is produced in the two upper block of the engine - Aggregator and Builder - based on a set of contextual rules driven by conditions on data already calculated.

The user's profile is normally created on the basis of the activities and interactions of users in the learning environment. Data in user profile is stored in form of logs and tracking data (even if some systems already memorized it in the learner model, at an higher level of aggregation and abstraction). Although many Learning Management Systems already provide the possibility to explore this user tracking data, in many cases the visual presentation of the information is not well suited to users' specific needs. Usually the exposed data is provided as a simple list of user activities, or access to course contents, without the possibility to explore it with an aggregated didactic view. In fact, this feature was originally thought for tecphnician in charge of solving technical issues, rather than for instructors or tutors interested in improving pedagogical aspects. In other cases, the presentation of data is limited to a data subset or is predefined by developers and fixed (Mazza & Milani 2004, Mazzola et al. 2010a). Notable exceptions in the field of OLM are OLMlets (Bull et al. 2009), in which the learner can choose between seven different representations. Nevertheless they normally relies on data coming from the single system they were developed with.

We provide an easy way to create effective graphical presentation of arbitrary data from different and heterogeneous sources. We propose a three-tier architecture composed by a data extractor, a data aggregator, and a builder (see Fig. 1). All the levels rely on a configuration file that the instructional designer can change or expand to create graphical indicators (in the form of widgets) of one or more interesting characteristics of the user profile. Our infrastructure can connect to any data source with different connection types (e.g. databases, Web services, connection bus), only by writing a small adapter. The output of our tool, as seen by a final user, is a flash based interface that represents, with one of the available graphical metaphors, the personal information aggregated according to the didactic model. We rely on a highly configurable infrastructure based on three layers: extraction, aggregation and widget creation. This schema follows a common data processing pattern: retrieve raw data, extract or derive , and present it in the most suitable way (Mazzola & Mazza 2009a, Mazzola & Mazza 2009b).

Mashup of data in web 2.0

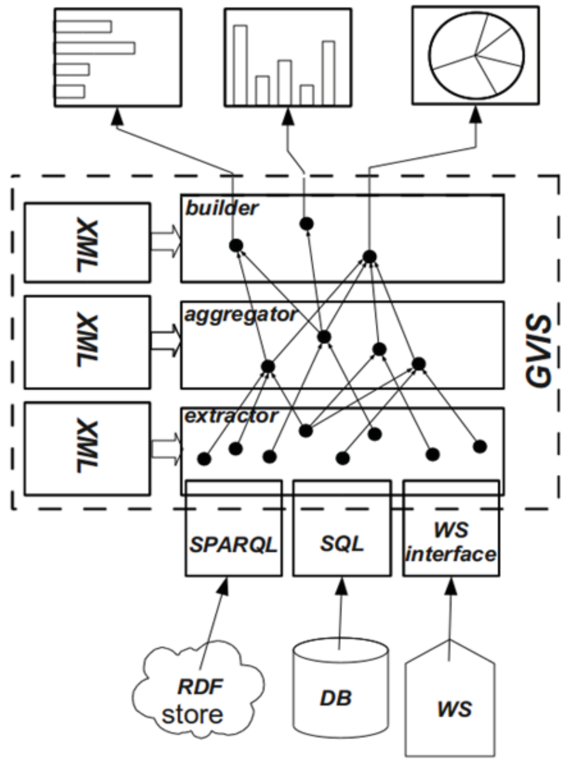

Fig. 2 The mashup architecture: the data path from heterogeneous sources to final smart graphical widgets.

Since Webservices and RDF specifications play a major role in integrating distributed services for Web 2.0, we developed an infrastructure able to take into account different kind of data sources by means of configuration profiles. The process of including facilities in a liquid and adaptable environment requires not only the availability of standardized way for namelessly connect the active component with the environment that will host it (known as the container), but also an effective method –like the JSON specification– for exchange data between different applications, services and data storage facilities. Fig. 2 shows the current infrastructure of GVIS, with a detail on the connection with heterogeneous sources. The system is flexible as the behavior of its main components can be easily changed by modifying their matching XML configuration files.

Even if the tool is designed for usage in the context of a distributed and heterogeneous environment, for a first test case, we applied this architecture to a single course deployed on our institutional LMS, that we will present in the next paragraph. Even if this case does not present distributed characteristics, it is important to consider the fact that the unique source of data is external to the tool. For the other aspect -the mashup of data-, it requires more than one source that are naturally related to the learning experience and is not easy to find them. As a proof of concept, we are starting using the same tool on a federation of different learning platforms in the context of the European project GRAPPLE. To test the practical feasibility of mashing up data we developed other experiences, relaying on completely different sources of data, like folksonomies and tracking logs from personal browsing history (Mazzola & al. 2010b).

A first application in TEL

In order to test the GVIS visualization module in a real environment, we created a testing platform with the Moodle LMS that includes a GVIS visualization module. We started without relying on the adaptive features. In this environment, the data source for GVIS is the log data created by the Moodle LMS and stored in its local database. The first step is the creation of meaningful didactic aggregators: in other terms, how to aggregate source data to helps learners and instructors in their activities. For achieving this goal, we adopted an iterative approach: the teacher indicates which kind of information is interesting for his activities and a Moodle expert proposes a possible way to collect the needed data. Subsequently, they inspect the results, decide what is to modify, and change it accordingly. At the end of this process we identified a set of six candidate indicators: two for the learners, very compact and seamlessly integrated into the usual Moodle users interface, and four for the tutor/teacher.

Fig. 3. The output of GVIS module inserted into a Moodle course.

This initial pilot phase was conducted applying the GVIS architecture to the course “Educational Communication and eLearning" held during the winter term of the academic year 2009-2010 by prof. Cantoni at Politecnico di Milano in Italy, in the MS degree in Engineering in Computer Science (see Fig. 3). Although in this test we did not aggregate data from different sources, our visualization infrastructure helped to aggregate data coming from different tables in the Moodle database, and to visually represent contextual information about the course and the learners. We developed a specific widget for representing the number of logins and the messages posted in a forum. An interesting outcomes was the graphical comparisons between the learner's specific information and the average value achieved by the class, which can work as reference for the self-monitoring process of the user's progress (Woolf 2009). We implemented functionalities for supporting learner: Fig. 3 shows the number of messages posted to the forums by the current user and the combination of accesses done to resources in the course by the student. The widgets also allow the comparison with the average of the class and, in the first case, also with the expected level by teacher. Both these pieces of information are considered important by the instructional designer who developed the online part of the course. Others widgets developed for this pilot study are the number of login to the course group by date and by student, and the number of forum post by date (grouped by the people that posted each day, the total number of post for day and for students).

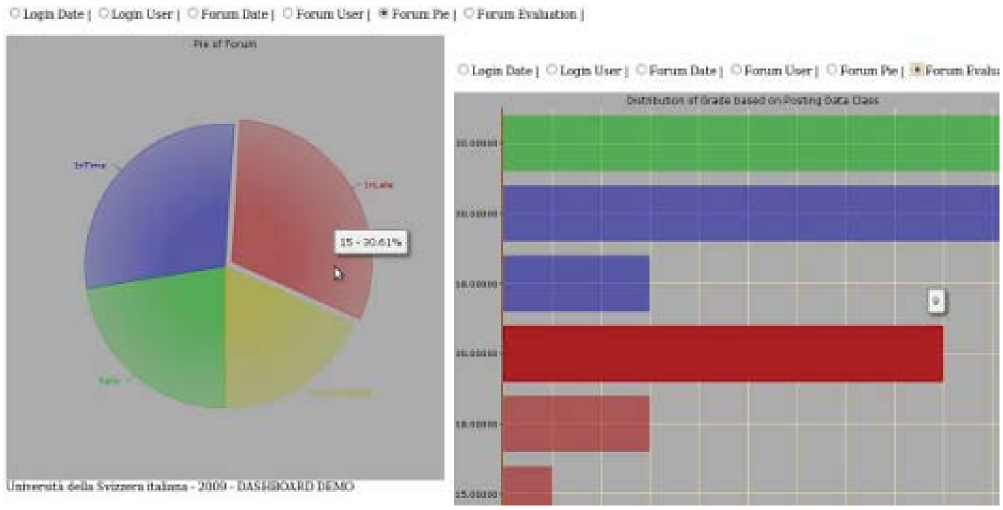

Fig. 4. The GVIS instance for teacher. A classification of the posts based on the submission date - Early, OnTime, Late and Uncompleted - is presented (on left), accompanied by the relative evaluation (on right).

Specifically (see Fig. 4) in the pie chart, groups of messages posted during the same range of dates, based on different deadlines, are depicted using different colors. Each slice of the pie represents a category based on the posting date with respect to the deadline set by the instructor; this can be: early, on time, late, or uncompleted. The size of the slice indicates the number of messages into each posting category. The instructor may give a grade to each message posted in the forums, depicted in the bar chart on the right. In this visualization, bars represent the distribution of grades given by instructor to postings, and the color of bars represent the posting specific category (early, on time, late, or uncompleted).

The user experience feedback

We run an initial evaluation that allowed the learners to provide feedback on their experience with the system. We devised a questionnaire with 16 question on a 5 point based Likert-scale. We submitted it at the end of the course as an online survey. The possible answer ranges from complete disagreement with the concept expressed (1 - SD) to complete agreement (5 - SA). The first interesting result was that 22 out of the 45 students answered in the timeframe allocated. The first impression seemed to be promising, even if we know that a real scenario is needed to confirm these initial results, due to the very limited impact of the current one (only one course during a semester).

For the moment we ignore the non-quantitative answers we received. After some introductory information about the aims and the scope of the survey, the questionnaire collected demographic data, such as the age, the gender and the role of the respondents. Unfortunately, probably due to the context of the course, we collected feedback only from male students. The average value for the age is 23.77 - with variation from a minimum of 22 to 32 as maximum - and all of them compiled the survey as learner. In Table 1 we report some analysis on the results: for easiness of interpretation we converted all the answer in a positive scale, complementing to the maximum each value obtained in the question marked by the asterisk.

We group variables for preforming some analysis on them; w ecalled the found sets C1, C2, C3 and C4, based on the percentage of positive versus negative answer provided and the standard distribution measure. In the first group, the most positive one, we collected the question USD1, USD2, and VBBI1 that are related to the easiness in understanding the visualizations, and to its capability of helping teacher in tailoring the teaching to individual learners.

	Cod	Question	nr	MIN	MAX	Mean	SD
C2	UST1	I find this visualization suitable for getting an overview on the current status in the learning process	22	2	5	3.86	0.7102
C2	UST2 *	I think the visualisation provides irrelevant information	22	2	5	3.86	1.0372
C1	USD1	It is easy to understand this visualisation	22	3	5	4.23	0.8125
C1	USD2 *	I find this visualisation unnecessarily complex	22	2	5	4.05	0.8439
C2	VBM1	I think this visualisation can help learners to reflect on their learning	21	2	5	3.71	0.7838
C3	VBM2 *	I think this visualisation will not significantly promote learners understanding and awareness of their learning progress	21	2	5	3.57	1.0282
C4	VBCL1	I think this visualisation is able to leverage mental workload	21	2	4	3.14	0.7270
C4	VBCL2 *	I think interpreting this visualisation would put additional cognitive effort on the learner	21	1	5	3.00	0.8891
C2	VBLE1 *	I think the use of this visualisations will not make a difference for learning performance	21	1	5	3.62	1.0713
C4	VBBP1 *	I think this visualisation would hinder collaboration among peers	21	2	5	3.10	1.1360
C2	VBBP2	I think this visualisation can help learners to better understand their learning through comparison with other learners	21	1	5	3.71	1.0556
C1	VBBI1	I think this visualisation would help instructors in tailoring their teaching to individual learners	21	2	5	3.90	0.8891
C3	VBBI2 *	I don't think that this visualisation can help teachers in better understanding their students' needs	21	1	5	3.57	1.3628

Table 1. The results of the survey: analysis

In set named C2 - composed of the question called UST1, UST2, VBM1, VBLE1, VBBP2 - the rate of positive answers is predominant but some learners reported to find that dimension are not so well suited. The topics explored in this group are related to the capability of offering a suitable overview of the current status, to the absence of irrelevant information, how much the visualization can support the users' reflection (also in comparison to peers) about their learning experience and the expected impact of the visualization on learning performances. The following two groups share the fact that negative answers arise to a significant rate, positioning them between the (possibly or quite surely) problematic aspects.

Cluster C3, that involve the questions VBM2 and VBBI2, are respectively related to the possibility to promote learners understanding and awareness of their learning progress, and to help teachers in better understanding the students' needs.

Furthermore, the most problematic set C4 is represented by the VBCL1, VBCL2 , and VBBP1 questions that investigate about the capability to leverage mental workload using GVIS, the additional cognitive effort imposed by the tool and the possibility that it prevents the possible collaboration among peers.

Conclusions

We presented an architecture to enable instructional designers to create graphical representations of one or more characteristics of the learner model. The semantic is expressed through a number of configuration files that drive the behavior of the component. We collected feedback from learner about their experiences and their feeling that the implemented tool is able to achieve the declared objectives. The evaluation of quantitative answers seems promising, even if some minor problematic aspects still exists. The contribution of this work is twofold: in the orchestration domain and also in the support for the sustainability of TEL solutions. For the former aspect two concurrent factors could positively influence the learner experience: the *adaptability* in order to reduce the cognitive overload required for the information interpretation and the *support* of metacognitive skills through the self-reflection processes by the learners. For the latter, instead, an enhancement of the *awareness* about the learning experience, through the visualization for presenting information, can play a major role in supporting the tutor duties and enabling teacher to improve the learning paths and activities, based on real usage data. We are aware of the possibility of achieve negative effects through the application of our infrastructure to TEL experiences, ranging from unexpected behaviors in groups with a low level of participation or riding phenomenons, to the impact of this data on learner with less active attitude to interact and collaborate till the informative overload for newly comers.

Acknowledgments

This work was partially funded by the EU FP7 project GRAPPLE, under the grant nr. IST-2007-215434 . We want to thank the numerous project partners for all the valuable ideas and contribution provided in this research, in particular we acknowledge University of Graz for providing us with the questionnaires and running some initial evaluations and Open University of Netherlands for the contribution in the design phase of the GVIS tool. For the test case we thanks a lot prof. Lorenzo Cantoni and all the students involved at Politecnico di Milano – Como campus in the evaluation phase.

References

(Bull 1997) Bull, S (1997). See Yourself Write: A Simple Student Model to Make Students Think. Proceedings of 6th International Conference on User Modeling, 1997.

(Bull et al. 2009) Bull, S., Gardner, P., Ahmad, N., Ting, J., and Clarke, B. (2009). \Use and Trust of Simple Independent Open Learner Models to Support Learning within and across Courses". LNCS, vol. 5535. Springer-Verlag, Berlin, Heidelberg, pp. 42-53.

(Bull & Kay 2008) Bull, S., Kay, J. (2008). Metacognition and Open Learner Models. Metacognition Workshop, Intelligent Tutoring Systems 2008.

(Chen 2009) Chen, C. (2009). Influence of Perceived Information Overload on Learning in Computer-Mediated Communication. Proceedings of the 8th international Conference on Advances in Web Based Learning, 2009.

(Dror et al. 2008) Dror, B., Nadine, M., Mike, B., (2008). Visualization and Analysis of Student Interaction in an Adaptive Exploratory Learning Environment. Proceedings 1st International Workshop in Intelligent Support for Exploratory Environments, EC-TEL'08

(Esposito et al. 2004) F. Esposito, F., Licchelli, O., Semeraro, G., (2004). Discovering Student Models in e-learning Systems. Journal of Universal Computer Science 10-1, pp. 47-57

(Heckmann & al. 2005) Heckmann, D., Schwartz, T., Brandherm, B., Krner , A., (2005). Decentralized User Modeling with UserML and GUMO. Proceedings of the Workshop on Decentralized, Agent Based and Social Approaches to User Modelling - DASUM 2005, Edinburgh, Scotland, pp. 61-65

(Mazza & Milani 2004) Mazza, R., Milani, C., (2004). GISMO: a Graphical Interactive Student Monitoring Tool for Course Management Systems. T.E.L.'04 Technology Enhanced Learning '04 International Conference. Milan, 18-19 November 2004.

(Mazzola & al. 2010a) Mazzola L., Nidola M., Milani C., & Mazza R. (2010**).** L'aggiornamento del tool di monitoraggio delle attività degli studenti: GISMO 2.0**.** Atti del convegno italiano MoodleMoot 2010 - Bari , Italia - ISBN 978-88-8459-162-3.

(Mazzola & al. 2010b) Mazzola L., Eynard D., & Mazza R. (2010). GVIS: a framework for graphical mashups of heterogeneous sources to support data interpretation. 3nd Conference on Human System Interactions, 2010. HSI '10, Rzeszow (PL), 578 – 584. ISBN: 978-1-4244-7560-5 DOI: 10.1109/HSI.2010.5514511

(Mazzola & Mazza 2009a) Mazzola, L. & Mazza, R. (2009). Toward Adaptive Presentations of Student Models in eLearning Environments. Frontiers in Artificial Intelligence and Applications, vol. 200. IOS Press, Amsterdam, The Netherlands, 761-762.

(Mazzola & Mazza 2009b) Mazzola, L. & Mazza, R. (2009). Supporting learners in Adaptive Learning Environments through the enhancement of the Student Model. Proceeding of Human Computer Interaction International 2009. San Diego, CA, 166-175.

(McCalla 2004) McCalla, G. (2004). "The Ecological Approach to the Design of E-Learning Environments: Purpose-based Capture and Use of Information About Learners." In Journal of Interactive Media in Education, 2004 (7).

(Schutz 2009) Schutz, D. (2009). Educational Intelligence: Bringing Schools into the 21st Century. Available online at SSRN: http://ssrn.com/abstract=1336243 Last accessed on : Oct, 10th, 2010.

(Van Der Sluijs & Houben 2006)Van Der Sluijs, K., Houben, G., (2006). A generic component for exchanging user models between web-based systems. International Journal of Continuing Engineering Education and Life-Long Learning 2006 - Vol. 16, No.1/2 pp. 64 - 76

(Van Der Sluijs & Hover 2009) Van Der Sluijs, K., & Hover, K. (2009). \Integrating Adaptive Functionality in a LMS." International Journal Of Emerging Technologies In Learning (IJET), 4(4). Retrieved December 31, 2009, from http://online-journals.org/i-jet/article/view/958

(Woolf 2009) Woolf B. (2009). Building Intelligent Interactive Tutors: Student–Centered Strategies for Revolutionizing E-Learning, Morgan-Kaufmann.