Extending Web-based Educational Systems with Personalised Support through User Centred Designed Recommendations along the e-Learning Life Cycle

Olga C. Santos a,*, Jesus G. Boticario a, Diana Pérez-Marín b

a aDeNu Research Group. Department of Artificial Intelligence. Computer Science School, UNED. Calle Juan del Rosal, 16. Madrid 28040. Spain. E-mail: ocsantos@dia.uned.es. E-mail: jgb@dia.uned.es
b Department of Computing Languages and Systems I. Escuela Técnica Superior de Ingeniería Informática. Universidad Rey Juan Carlos. Avenida Tulipán s/n, Móstoles 28933, Madrid. Spain. E-mail: diana.perez@urjc.es

*Author to whom correspondence should be addressed. Phone: +34 91 398 93 88;

Abstract: In this paper we address an open key issue during the development of web-based educational systems. In particular, we provide an educational-oriented approach for building personalised e-learning environments that focuses on putting the learners’ needs in the centre of the development process. Our approach proposes user centred design methodologies involving interdisciplinary teams of software developers and domain experts. It is illustrated in an adaptive e-learning system, where a MOOC (Massive Open Online Course) was taken by nearly 400 learners. In particular, we report where user centred design methods can be applied along the e-learning life cycle to designing and evaluating personalisation support through recommendations in learning management systems.

Keywords: e-learning life cycle, personalisation, adaptive educational systems, educational recommender systems, user centred design

1. Introduction

As discussed in this paper, User Centred Design (UCD) approaches can be used to enrich the personalisation capabilities of web-based educational systems with adaptive navigation support through recommendations. From our experience, personalisation cannot be designed in advance as a plug-in in the e-learning environment, but it has to be constructed in a process that considers and involves learners’ needs from the outset and during the whole e-learning life cycle (eLLC). To cope with this need, we propose the adoption of a UCD approach during the development of adaptive e-learning systems, involving interdisciplinary teams of software developers and domain experts. As a result, we have come up with a solution for building personalised e-learning
environments that has shown a better support for teaching and learning activities by putting the user in the centre of the development process. This is a long-standing challenge during the development of services and functionalities for new e-learning solutions and is becoming critical when developing Massive Open Online Courses (MOOCs) that are to cope with individual learners’ needs.

In order to address this challenge, at aDeNu research group we have been working for more than a decade on how to take adaptation along the eLLC by designing and delivering educational-oriented recommendations. In our methodology, four phases are handled in this cycle: i) design in advance the learning experience, ii) administrate the environment where the learning experience is to be carried out, iii) interact (learners and educators) with the e-learning services, and iv) feed the results on the learning experience back to the course author. Here, adaptation is delivered by personalised recommendations through a Semantic Educational Recommender System (SERS). In previous works we have partially presented the basis of this approach, but we have not reported how each of these parts can be combined along the development cycle of the e-learning system to provide the required functionality. Neither, we have reported the results of a large-scale evaluation covering all the cycle stages to validate the approach.

These partial advances have been achieved within some of the research projects developed at aDeNu. In particular, the grounds for the definition of the stages in the eLLC, the need to involve all the actors of the learning process and the idea of combining design and runtime adaptations making a pervasive use of educational standards were settled in the aLFanet project (IST-2001-33288) [1, 2]. Following these, in the ALPE project (eTEN-2005-029328) usability and accessibility issues were included in the eLLC [3]. Finally, in the EU4ALL project (IST-2006-034778) we deepened on the previous ideas and defined an open standard-based service-oriented architecture to extend the personalisation capabilities of learning management systems (LMS) in an inclusive way [4, 5], set the foundations for the UCD TORMES methodology to help educators elicit and design educational-oriented recommendations [6, 7], proposed the recommendation model that bridges the gap between the educators’ descriptions and the software components’ information exchange [8, 9] and drafted the system requirements to instantiate elicited recommendations within the LMS [10]. Based on these works, an approach to extend e-learning systems with personalised support with user centred designed recommendations along the eLLC is proposed.

This approach has been tested in a real world e-learning scenario by deploying a MOOC through Willow web-based educational system enriched with user centred designed educational-oriented recommendations. Willow is a computer-assisted assessment system of free-text answers successfully used in blended learning [11]. The integration of the SERS approach in Willow was outlined in [12]. As addition to the work reported there, this paper provides details on the user centred development and evaluation process carried out along the eLLC by an interdisciplinary team of software developers and domain experts to extend Willow’s adaptive capabilities in order to provide learners with the required adaptive navigation support. In particular, in this paper we provide some practical guidelines that focus on reporting where UCD methods can be applied along the eLLC to designing and evaluating personalisation support in e-learning systems through recommendations.
The paper is structured as follows. First, some background and motivation is introduced. Then, our approach to consider personalisation along the eLLC is presented. After that, we report its application to produce user centred adaptive support in the aforementioned MOOC. Finally, some discussion, conclusions and future work are outlined.

2. Background and Motivation

Adaptive e-learning systems have undergone considerable changes over the last decade [13]. There have been research prototypes [14] but still personalisation in e-learning is a long-standing open issue [15] and there is not currently any system that supports full adaptivity [16, 17]. A straightforward way to provide adaptation in LMS is via adaptive strategies like adaptive ordering, link hiding, and adaptive link annotation with the aim to tailor the learning experience to each individual learning need [18]. In this context, LMS can be seen as a large space of educational information sources and interaction items, which include not only the educational contents, but also contributions produced and shared by members in the course (e.g. forum posts, external links, etc.). These potential navigation paths involve an increasing number of possible actions to take in the course space [19]. However, although LMS are holding large volumes of interaction data, they do not integrate any data mining support to evaluate the course space structure and its effectiveness on the learning process, aimed to develop smarter scenarios that enhance the learning experience [20]. For this reason, data mining services that can be integrated into LMS to support educators (non-experts in data mining) in the data mining process are being proposed [21].

Recommender Systems (RS) are also being integrated into LMS to reduce the existing information overload and extend their capabilities with adaptive navigation support [22]. This integration can be carried out following a modular approach based on building components that support different services that can be integrated into different platforms, as devised for the next generation of LMS [23]. More specifically, service-oriented architectures can contribute to improving the personalisation capabilities of LMS [24, 25], for instance, by integrating external services that manage the learner model upon which the adaptation can take place. In the same vein, standardisation approaches towards the Web 3.0 are being carried out, aiming at enabling interoperability with existing services and developments at different institutions [26]. In this context, one of the most holistic standardisation approaches in education is the set of specifications developed by IMS [27], which include, among others, the Learning Design specification. IMS Learning Design offers the explicit semantic description that can be managed by educators in providing adaptive learning paths [28]. Similar information is also relevant for adaptation purposes in educational RS (ERS) [29].

A relevant issue from the user viewpoint is the concept of usability, defined the standard ISO 9241-11 as “the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use” [30]. In e-learning, the concept of usability has been discussed for years [31] and is considered a valuable quality criterion for e-learning systems [18], a basic condition for their effectiveness [32] and a key factor in the user’s ability to acquire knowledge in a satisfactory manner [33]. The usability of the learning experience should also consider the pedagogical value of the e-learning designs [34], since an e-learning application may be usable but not in the pedagogical sense, and vice versa [35]. Furthermore, the usability design principle of consistency, which states that users should not have to wonder whether different words, situations, or actions mean the same thing [36], has to
be taken into account when RS are integrated into LMS to make this integration as less intrusive as possible, which is a desired goal in RS [37].

When ERS are designed, as for any RS, the focus should be put on understanding the users’ needs [38]. ERS can cope with a wide range of situations that may arise over the learning process and guide the learner under different circumstances. For instance, as discussed in [7], they can be used not only to point to relevant contributions from other users, but also to foster their active involvement in the learning process suggesting actions to carry out with the available learning resources. Nevertheless, several recent works that compile the state of the art on ERS (e.g. [7, 39, 40, 41, 42, 43, 44]) show that to date most of the recommendations have focused on finding suitable materials as an alternative to relying solely on others’ opinions, as originally proposed [45]. However, a recommendation in the e-learning context could be much more interactive [19].

Therefore, in order to understand users’ needs in the personalisation process, and specially in the recommendation process, UCD approaches, which develop systems that suit users’ needs [47], should be followed. In particular, UCD creates experiences for people with their needs in mind and is mainly focused on achieving active involvement of users in designing development tasks in order to improve the understanding of user and task requirements [48]. Thus UCD should be central to the development of any adaptive e-learning system inasmuch as learning is a personalised and evolving process that must be focused on the learner [49]. Regrettably though, UCD is not usually considered when developing adaptive e-learning systems [50]. In fact, none of the 59 systems analysed (see [19] and references from [51] to [108]), has been designed following UCD. If considered, UCD grounds the process in information about the people who will use the system and supports the entire development process with user centred activities in order to create applications which are easy to use and are of added value to the intended users [109]. When involving users in UCD, people with disabilities are also to be included to properly consider accessibility requirements, so the resulting developments are universally usable [110].

Testing with real users is important for the RS design, and criteria to capture the user perception of the recommendations have to be explored [111]. Evaluation approaches reported in the literature of ERS show that the field is in an early stage. This lack of evaluation approaches is inherent to adaptive systems in general [112]. However, some generic frameworks for interactive adaptive systems have being defined following a layered evaluation approach, where adaptation is broken down into its constituents (called layers) and each of these layers is evaluated separately [113]. In addition to this, a couple of user centric evaluation frameworks for RS have also been proposed [114, 115], but none of them discusses their applicability for educational scenarios. And this is not a trivial issue, since the educational domain differs largely from other recommendation scenarios in goals, user models and environmental conditions [116].

When evaluations on ERS are reported, they usually focus on evaluating the recommendation algorithm’s performance with end-users, but educational-oriented evaluations that measure their effect on the learners are still scarce. This suggests that specific approaches to measure the effect on the learning process are still unexplored and confirms others conclusions [116, 46]. From the review carried out to the aforementioned 59 systems, we found 16 publications that do not report any evaluation at all [19, 52, 53, 55, 62, 63, 65, 82, 88, 90, 94, 95, 98, 101, 107, 108], although in most
cases the evaluation was mentioned as future work. Following the categorization approach proposed in [46] for the other 43 works, the evaluation focus of half of them (27) was put only on evaluating the algorithms [54, 56, 57, 59, 60, 61, 64, 66, 68, 69, 72, 73, 76, 78, 80, 83, 84, 86, 89, 91, 93, 96, 99, 100, 102, 103, 105]. The system usage was evaluated in 8 works [51, 71, 74, 77, 79, 81, 92, 104], half of them focused on usability issues [71, 79, 81, 104]. Only 6 of the systems evaluated the users’ perceptions on the recommendations [67, 71, 75, 79, 85, 87] and 7 explicitly evaluated the learning performance, mainly in terms of knowledge gain [58, 70, 71, 85, 87, 97, 106]. Regarding the evaluation role, in most cases (25) evaluations were done on end users [51, 57, 58, 60, 61, 67, 68, 70, 71, 73, 74, 75, 76, 77, 79, 81, 85, 86, 87, 92, 93, 97, 99, 104, 106]. In the rest, systems designers [56, 64, 78, 80, 84, 89, 102, 103, 105] or simulated learners [66, 72, 91, 96] were used.

Thus, for the many different problems involved, a hot research issue in RS has been the evaluation approach. From the system’s point of view, several evaluation metrics that deal with recommendations accuracy and error (broadly defined) have been proposed. They focus on evaluating algorithms off-line to avoid presenting recommendations that might not be accurate [117]. However, when this goal is achieved and the recommendations produced are reasonably accurate, then these authors suggest putting the focus on evaluating the quality of recommendations themselves. This involves understanding how they are perceived by the user and how personalisation and recommendation impact on the user during and after interaction [118]. In fact, one of the goals of an evaluation process in RS is to show the usefulness of the system and to find out if the system is valuable to the users [38]. This research problem is even more relevant in the educational domain, where the primary users of a learning system are the learners and recommenders are expected to improve their learning somehow [116].

Thus, there have been two different and complementary approaches that can be used to evaluate RS [119]: 1) off-line evaluation (which evaluates the performance of the recommendation mechanism on existing datasets), and 2) on-line evaluation (which evaluates the recommender performance on the effects on users of a running RS). To carry out off-line evaluations, metrics from the machine learning and information retrieval fields have been borrowed. These metrics are based on statistical measures and focus on accuracy, measuring empirically how close the recommender’s predicted value differs from the true value. However, off-line analysis has limitations [117] even to measure the true accuracy of recommendations [118]. Several studies have shown that user satisfaction does not always correlate with high RS accuracy [120, 121]. Thus, accuracy measures are not useful to measure the user’s perception [122]. For this, online evaluations involving users on a running system are required.

In this way, the success of RS in influencing users can be measured through the change in user behaviour [117]. In fact, RS should be evaluated in terms of user experience [123], which, according to the standard ISO 9241-210 is defined as “a person’s perceptions and responses that result from the use or anticipated use of a product, system or service” [124]. In e-learning, the presumption is that the more learners perceive usefulness and ease of use, the more positive their attitudes are toward e-learning, consequently improving learning experience and satisfaction [125]. In any case, users are also to be involved in evaluating the user experience.
Thus, in order to enrich the personalisation capabilities of web-based educational systems, users are to be involved in the design and evaluation of the adaptation provided. Moreover, taken advantage of the service-oriented architecture approach and available standards and specifications, external recommendation services can be integrated in existing LMS to provide adaptive navigation support. In this context, we have proposed a UCD-based approach that covers the four phases of the eLLC.

3. Personalised Support along the e-Learning Life Cycle

To cope with well-known topics in adaptive education such as adaptive navigation support, a personalised e-learning process cannot be solely managed in terms of design-based approaches (e.g., instructional or learning design) or run-time solutions (e.g., educational data mining) [1]. Given the complexity of the learning context, it is not possible to design in advance the most appropriate navigation path for each learner in each situation as instructional design theories implicitly assume [126]. A personalised navigation has to take also into account the features of the actual learner [127] and her current and past interactions [68] along with the education issues that have been involved. In this sense, regarding educational issues, some relation has been found between users’ cognitive style –i.e. the approach to organizing and representing information– and navigation behaviour [128], showing that learning patterns have effect on learning performance [129]. However, mining interaction data does not cover all the educational issues that can be addressed from the instructional design approach. In fact, to identify high quality instructional designs best practices from educators are to be gathered [130], and approaches to support them in capturing their view of the educational process are being proposed [131].

In order to take advantage of the educators’ expertise while learners’ needs are taken into account, both main actors of the learning process (i.e. learners and educators) are to be involved and supported during the eLLC. The approach presented here is based on aLFanet’s idea of combining design and runtime adaptations [1] and integrates ALPE’s extension in the eLLC to consider the usability and accessibility needs [3]. As a result, user centred adaptations are applied along the eLLC covering the following phases (as shown in Figure 1): 1) design: deals with the preparation in advance of the learning experience, packaging the instructional design produced in a way that can be imported in an LMS, and considering usability and accessibility requirements (via requirements gathering and elicitation methods) as well as evaluation results from the feedback phase of a previous iteration cycle (if any); 2) publication: imports the instructional design and administers the environment where the learning experience is to be carried out (i.e. the learning space), performing usability and accessibility expert reviews to detect problems appearing when the course is imported into the LMS and the corresponding learning services are configured; 3) use: focuses on the usage of the LMS services by learners and educators in a usable and accessible way; to this empirical testing with user-based methods (e.g. direct observation using thinking aloud and contextual inquiry techniques) and task-based scenarios can be applied to track the learners’ behaviour and learn about their interactions; and 4) auditing: provides feedback to the course author on the learners’ experience; to this experts, users’ evaluations and automatic processing of interactions are to be analysed with usability metrics and reported back to the design phase of the next iteration.
To address adaptive learning in itself, these four phases should be focused on i) the learners’ needs, ii) the interaction preferences in accessing the course contents (including the accessibility issues involved), and iii) the educational support required when carrying out the course activities. Thus, adaptation is a process that influences the eLLC and draws on an information management process among the phases, which can be semantically modelled by specifications and standards, as discussed elsewhere [4]. Figure 1 shows the main goals of each of the eLLC phases and the resulting outcomes passed from one to the other. Moreover, an additional phase (named system in-house debugging) can be carried out before the design phase when a new LMS is installed or a new version is upgraded. The goal is to carry out a usability heuristic evaluation and the analysis of the accessibility guidelines on the LMS to be used at the institution in order to report required improvements and ask developers to make corrective actions if needed (and possible).

In order to provide personalised dynamic support to learners when taking a course in an LMS, in the aLFanet project an ERS was proposed and developed. The goal was to provide runtime adaptation to cope with the learning needs and unpredictable situations that learners come across while interacting with the system through learning paths specified in IMS Learning Design [1]. In particular, two recurrent pedagogical situations in on-line courses were addressed (i.e., lack of knowledge and high interest). In these situations, recommendations were offered when the user context matched the conditions defined. From these and other related experiences carried out in several projects were the aDeNu research group has been involved, this ERS approach evolved towards the concept of Semantic ERS (SERS), described in detailed elsewhere [9]. Very briefly, SERS follow the service-oriented approach to extend LMS with adaptive navigation support and manage recommendations described in terms of a semantic recommendation model that specifies (using standards and specifications when possible) ‘what’ should be recommended, ‘where’ and ‘how’ the recommendation should be communicated, to ‘whom’ and ‘when’ the recommendation should be offered, ‘why’ the recommendation is delivered, and ‘which’ are the recommendation features.

From the above follows that when extending the adaptive support of LMS with recommendations derived from the SERS approach, the eLLC process and its phases
have to be considered. In particular, first, if not done before, the system in-house debugging is needed to evaluate not only the usability and accessibility of the LMS, but also the availability of the functionality required to publish and deliver the recommendations. Second, at the design phase, which deals with the preparation of the materials for the learning process, UCD methods (such as stakeholders meetings, user surveys, interviews, user observations, scenarios of use, card sorting, focus groups, Wizard of Oz) are to be used to elicit educational sound recommendations from experts in on-line teaching and validate them with learners [7]. The resulting recommendations are to be described in terms of a semantic recommendations model so they can be published in the SERS that provides the recommendation service to the LMS [8]. Third, the publication phase, which deals with the preparation of the environment for the learning process, involves instantiating the previously designed recommendations into the SERS so they can be delivered in the learning space when appropriate [9]. Fourth, the use phase, which supports the learners’ interactions with contents and services in the LMS, is in charge of offering the learners personalised recommendations whose semantic description match the runtime context [8]. In this phase, both quantitative (from data logs) and qualitative (from pre and post test, and satisfaction questionnaires) data are obtained. Finally, on fifth place, the auditing phase, which deals with assessing the learning experience by analysing previously gathered interactions to evaluate recommendations impact on learners [12]. These results are fed back to the design phase to adjust future iterations and if needed, modify recommendations design.

In this context, TORMES methodology has been proposed to support educators in eliciting educational-oriented recommendations [6]. In TORMES, data mining techniques can be used to enrich the knowledge obtained with UCD methods, as described in [7], where the SERS approach was deployed in dotLRN LMS. In that work, association rules and decision trees were used to identify reasonable indicators for having course success (e.g., filling in the user profile, using communication services such as forums or chat and doing self-assessment questionnaires). Moreover, as discussed there, clustering techniques can also be used to refine the applicability conditions for recommendation delivery (e.g., number of sessions that identify users with low participation, number of contributions in forums when users are very active).

Furthermore, TORMES can be extended with the layered evaluation approach to guide the formative evaluation of the adaptive features design (i.e. the recommendations). The five typical layers of the layered evaluation approach (collection of input data; interpretation of the collected data; modelling of the current state of the world; deciding upon adaptation; and applying adaptation) can be mapped into both the eLLC and the UCD cycle as defined in the standard ISO 9241-210 [124]. To the best of our knowledge and from recent literature review [132], TORMES can be considered the first attempt to apply the layered evaluation approach to the development of RS.

In order to facilitate the application of this approach during the design of web-based educational systems, the proposed steps to follow are outlined in Table 1. For each of the eLLC phases, it identifies its goal, the tasks to be carried out by each of the involved actors and the expected input and output. Thus, Table 1 can serve as practical guidelines that other researchers can follow to extend their LMS with adaptive navigation support.

<table>
<thead>
<tr>
<th>eLLC phases</th>
<th>Goal</th>
<th>Tasks</th>
<th>Actors</th>
<th>Input</th>
<th>Output</th>
</tr>
</thead>
<tbody>
<tr>
<td>System in-</td>
<td>Debug the</td>
<td>(1) Detect usability and accessibility</td>
<td>Usability</td>
<td>LMS installed</td>
<td>Debugged LMS</td>
</tr>
<tr>
<td>eLLC phases</td>
<td>Goal</td>
<td>Tasks</td>
<td>Actors</td>
<td>Input</td>
<td>Output</td>
</tr>
<tr>
<td>-------------</td>
<td>------</td>
<td>-------</td>
<td>--------</td>
<td>-------</td>
<td>--------</td>
</tr>
<tr>
<td><strong>house debugging</strong> (only when a new version of the platform is upgraded or a new platform is installed)</td>
<td>LMS to be enriched with adaptive navigation support</td>
<td>problems in the LMS (usability heuristics and accessibility guidelines evaluation)</td>
<td>and accessibility expert (Task 1)</td>
<td>in the educational institution</td>
<td>with usability and accessibility issues solved and added functionality for recommendations delivery</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(2) Check the availability of the functionality required to publish and deliver the recommendations</td>
<td>Software developer (Tasks 2,3)</td>
<td>Potential LMS objects and actions to be recommended</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(3) Modify the LMS to correct issues identified</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Design</strong></td>
<td>Prepare in advance the learning experience</td>
<td>(1) Design learning flow and prepare course materials</td>
<td>Educator (Tasks 1,2,3,4)</td>
<td>Feedback from previous auditing phase (if available)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(2) Identify user requirements with UCD gathering and elicitation methods (TORMES: stakeholders meetings, user surveys, interviews, user observations, scenarios of use, card sorting, focus groups, Wizard of Oz) + feedback from previous cycle auditing phase, including data mined from past course interactions</td>
<td>Learner (Task 4)</td>
<td>User requirements (learners’ needs, interaction preferences, educational support required)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(3) Describe recommendations elicited in terms of the semantic recommendations model</td>
<td>Usability and accessibility expert (Tasks 1, 2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(4) Validate recommendations</td>
<td>Knowledge engineer (Task 2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(5) Evaluate usability and accessibility of course design (usability heuristics and accessibility guidelines evaluation)</td>
<td>Educational support officer (Task 6)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(6) Package the instructional design produced</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Publication</strong></td>
<td>Prepare the learning space</td>
<td>(1) Import the packaged course design into the LMS</td>
<td>Educational support officer (Tasks 1,2,3)</td>
<td>Packaged course design</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(2) Administrate the environment (add users, configure LMS services)</td>
<td>Usability and accessibility expert (Task 4)</td>
<td>Semantically modelled recommendations</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(3) Add designed recommendations into the SERS</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(4) Evaluate usability (heuristics) and accessibility (guidelines) of the learning space</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Added users, imported contents, created recommendations and configured services in the learning space</td>
<td></td>
</tr>
<tr>
<td><strong>Use</strong></td>
<td>Support learners’ interaction with contents and services</td>
<td>(1) Use the learning services and receive personalised recommendations when needed</td>
<td>Learner (Task 1)</td>
<td>Learning space ready to be used (users, contents, services and recommendations)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(2) Track learners’ behaviour and learn about their interactions with empirical testing: e.g., user-based methods (e.g. direct observation using thinking aloud and contextual inquiry) and task-based scenarios</td>
<td>Usability and accessibility expert (Task 2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Auditing</strong></td>
<td>Assess the learning experience results to adjust future</td>
<td>(1) Analyse interactions: quantitative (from data logs) and qualitative (from questionnaires)</td>
<td>Educator (Tasks 1,2,3)</td>
<td>Interaction data in the learning space (both quantitative data from logs and qualitative data from pre test, post test, and satisfaction questionnaires) are obtained</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(2) Evaluate with usability metrics recommendations impact on learners</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Analysis of learning experience</td>
<td></td>
</tr>
</tbody>
</table>
As shown in Table 1, each phase receives an input, involves specific actors who are to carry out certain tasks to achieve a given goal and produce the corresponding output. The approach proposes design methodologies involving interdisciplinary teams of software developers and domain experts (i.e. educators, usability and accessibility experts, knowledge engineer, educational support officers), as well as the end-user, i.e., the learner. The later guides the design process from the beginning, as learners’ requirements are taken as input in the design phase. In that phase, potential learners are also involved in evaluating the recommendations elicited. In any case, all efforts are led to support their interactions in the use phase. Software developers might be required in the system in-house debugging phase to identify (and solve) limitations in the system functionality regarding the potential recommendation goals, and improve the usability and accessibility of the LMS, provided that this is feasible (e.g., the institution uses open software). Usability and accessibility experts are needed to evaluate the LMS to be used in the system in-house debugging phase, the contents in the design phase, the configuration of the learning environment in the publication phase, apply gathering and elicitation methods in the design phase and carry out the empirical testing in the use phase. A knowledge engineer can extract implicit knowledge from past courses interactions with data mining techniques. In turn, the educator is in charge of producing the instructional design of the course and eliciting the recommendations in the design phase with the help of TORMES methodology, as well as of analysing the learning experience in the auditing phase. The educational institution can provide an educational support officer to help the educator package the course design in the design phase, and import it in the LMS and configure the learning space in the publication phase. If this role is not available, the educator herself has to carry out these tasks.

The application of UCD methods is done along all the eLLC phases. From our experience, we have selected those methods that can provide information for the development process. Nevertheless, this proposal can be extended with other usability methods supporting UCD from those defined in the standard ISO/TR 16982:2002 [133]. In our proposal (as compiled in Table 1), heuristic evaluation is done in the system in-house debugging, design and publication phases. Gathering and elicitation methods, such as stakeholders meetings, user surveys, interviews, user observations, scenarios of use, card sorting, focus groups, and Wizard of Oz are used in the design phase as defined in TORMES elicitation methodology [6]. Empirical testing with user-based methods (e.g. direct observation using thinking aloud and contextual inquiry techniques) and task analysis are helpful in the use phase to understand the learner behaviour. Finally, usability metrics and qualitative feedback (e.g. from user surveys) can provide insight of the learning experience. In this way, we have proposed where UCD methods can be applied along the eLLC to designing and evaluating personalisation support through recommendations in LMS. Thereby, we provide some guidelines for building personalised educational systems that support teaching and learning activities and thus, we aim to address one of the key open issues during the
development of new e-learning systems that are to cope with the individual needs of a large number of learners, such as in current MOOC scenarios.

Next, we present how we have applied this approach to provide recommendations in a MOOC offered in Willow web-based educational system.

4. The e-Learning Life Cycle in a Real World Scenario
We have researched the integration of SERS into LMS to enrich the latter’s adaptive capabilities with an adaptive navigation support that takes into account the educational issues involved [9]. The review of the state of the art shows that most of the approaches that have been followed are similar to those used in other domains without taking advantage of the particularities of the educational issues involved [7]. For instance, recommendations in the educational domain should not be guided only by the learners’ preferences but also educational criteria should be considered [67].

In this context, our approach to consider user needs along the eLLC has shown its value, as reported in this section. In particular, a free-text adaptive computer assisted assessment system called Willow, which asks questions to learners about course concepts and provides feedback to them on their responses using the dialogue metaphor (i.e. an avatar “talks” to the learner -who in turn is represented by another avatar- on behalf of the system), has been extended with recommendations [12]. The adaptation goal was to provide educational-oriented recommendations that go beyond recommending courses or learning content and involve meta-cognitive issues. The course was delivered in a real world scenario completely on-line without any face-to-face interaction following the MOOC approach [134], and thus, aimed at large-scale participation and open access via the web. Two educators were involved in the experience, being supported along the eLLC in the elicitation, design and evaluation of the educational-oriented recommendations by a software developer, a usability and accessibility expert, a knowledge engineer and an educational support officer.

Next we comment on the application of the proposed approach to Willow throughout all the eLLC phases involved. The guidelines compiled in Table 1 have been followed.

4.1 System in-house debugging
Previous to the recommendations design, the functionality, usability and accessibility of the system where the MOOC was to be offered (i.e. Willow) was evaluated by an expert in these tasks. Some lacks were found regarding the functionality, as Willow was designed for blended learning. Thus, corrective actions were done by a software developer. In particular, Willow was extended with typical LMS tools (file-storage, forums, etc.). Moreover, in order to be able to deliver recommendations, and following the SERS approach [9], a service-oriented architecture was considered to extend Willow with the required adaptive navigation support.

Willow is an adaptive computer assisted assessment system, which provides user modelling capabilities. The user model consists of a hierarchy of concepts (i.e., basic, topic and area-of-knowledge concepts) and their relationships, defining a learner conceptual model. Whenever a learner answers a question in Willow, the system compares the use of the concept by the learner with the answer provided by the educator, and assigns a confidence value between 0 and 1. This value reflects how well the learner knows each particular concept according to a set of heuristics explained in
which combine statistical techniques (focused on processing the style of the answer, dealing with synonyms and word sense disambiguation) and latent semantic analysis (focused on processing the answer content to find similarities in semantics among learner’ and educator’ answers). With this information, Willow can build both the conceptual model of each learner as well as the conceptual model of the whole class. Both models are represented as concept maps (using a size and colour scheme to give information about the type of concept and its confidence value). In this way, Willow tracks the learner’s progress in the course regarding each of the concepts to be learned.

As the user model is already integrated into Willow, no external user model component was needed. However, in order to provide adaptive navigation support through recommendations, Willow’s user model was enriched with the processing of additional usage data: i) number of visits to each of the course pages, such as the page with the content lessons to study, the page with the questions to answer, the page with the concepts definition, the page to select the avatars, etc., ii) number of contributions in each of the platform services such as the forums, the questionnaire service, etc., iii) number of questions answered per lesson, iv) number of correct questions per lesson, v) number of sessions (i.e. periods of time that the learner interacts with the system) spent in the course, vi) recommendations offered, vii) recommendations followed, and viii) feedback given by the learner for the recommendations followed to report whether the recommendation was found useful, not needed or better to be offered in another moment. In this way, the following information about the learner behaviour can be obtained for each learner: course pages visited, time spent in sessions, questions answered, confidence values per concepts, contributions and visits in learning services, recommendations offered and followed, and feedback on the recommendation.

Regarding the recommendations service, a component was developed to provide this functionality. This service offers an XML-based interface description language following the W3C WSDL specification that provides a machine-readable description of the requests that the service can serve. In this way, it describes how the service can be called, what parameters it expects, and what data structures it returns. In particular, the functionality offered to the LMS consists in selecting the appropriate recommendations for a given learner taking into account her current learning context (getRecsForUser), inform the recommendation service that a recommendation has been followed (setRecClicked) and report that the learner has provided feedback to it (setRecFeedback). In turn, the recommendation service needs to get information from the user model (getUMInfo). Thus, to support interoperability, Willow had to implement calls for the required requests (getRecsForUser, setRecClicked and setRecFeedback) and allowed to be called by the recommendation service with getUMInfo. Recommendations logic was defined in terms of rules implemented with if-then sentences coding the applicability conditions defined by the educators in the design phase (see Table 3). Each if-then rule can have a variable number of applicability conditions, as follows: “IF applicability_condition_1 AND applicability_condition_2 AND ... AND applicability-condition_N THEN Ri”, where N is a natural number and Ri is the identifier of the recommendation to be offered. Each applicability condition is defined in terms of the content to be evaluated, the operator used for the evaluation and the value to which it has to be compared. Rules defined are listed in Section 4.3.

A three-layer architecture has been followed to implement this approach, where several services (i.e. services layer) are offered to the user through the Graphical User Interface
(GUI) (i.e. presentation layer). These services are the following: 1) User management: registers the users in the system, controls the access rights, manages the registration data, etc.; 2) Avatars management: allows the learner to select the avatars for the system and for herself; 3) Forum management: includes creation of forums, creation of threads, replies to threads; 4) Assessment management: manages the assessment functionality, that is, the publication of course pre-tests and post-tests as well as satisfaction questionnaires; 5) Materials management: manages uploading and downloading course materials; 6) Concept reviewing process: deals with the core functionality of Willow; 7) Observation of actions: tracks all actions carried out in the system, both active (i.e. contributions) and passive (i.e. visits); and 8) Recommendation process: selects the appropriate recommendations to be delivered to the current learner. In turn, users features and interactions, recommendations delivery data and course conceptual structured as well as the system web pages are managed by the persistence layer. Moreover, in order to show the recommendations to the learner, Willow GUI was extended following the dialogue metaphor, as shown in Figure 2. As a result, now the system suggests a set of recommendations (top dialogue) and the learner can provide feedback on the last recommendation followed (bottom dialogue) by selecting, if desired, one of the three clickable options available (useful, not needed, or better in another moment), as commented above.

![Figure 2. Integration of recommendations in Willow complying with the dialogue metaphor](image.png)

In this way, recommendations have been integrated in Willow in a non-intrusive way. They are delivered following the same interaction approach used by the system.

### 4.2 Design phase

In this phase, the learning experience was prepared in advance, taking into account the user requirements. Willow has been used by university learners both in engineering and non-engineering degrees since 2005 [11]. In all the experiences previous to the one reported here, the course contents were given in face-to-face sessions, following a blended-learning approach. In this context, two educators, with experience in using Willow in blended learning and teaching undergraduate learners of Pedagogy, had perceived, from their contact with these learners, the learners’ need to enrich their skills on browsing the Web to support their teaching practise. As a result, they decided to be involved in the development of a MOOC in Willow on ‘Search strategies in the Web with Educational Goals’ (EBIFE as abbreviated in Spanish), which was meant to be open to any interested teacher. They designed the learning flow and prepared the corresponding materials, which were checked by a usability and accessibility expert. This course was divided into two modules, M1 entitled ‘Sources of educational
resources in the Web’ and M2 entitled ‘Search strategies’. Contents were packaged by the educational support officer to facilitate their later deployment in any LMS. Recommendations were designed to foster educational issues while guiding learners’ interactions when using Willow by themselves, addressing those navigational issues that up to that moment were solved by the educators in the face-to-face introductory session.

To identify valuable educational-oriented recommendations, and in order to address the aforementioned related issues, TORMES UCD-based approach (which adapts the four activities of ISO 9241-210) was used to help the educators in eliciting valuable recommendation opportunities [6]. First, the context of use was gathered in the so called ISO 9241-210 activity. Here, the usability and accessibility expert interviewed the educators and asked them about the problems found by their learners in the past blended learning context, as well as about potential problems that could take place in an e-learning context from their perspective. One of the problems identified was that learners did not use Willow to review the course contents, even though they had entered in the system. Observations and data gathered by the educators from interaction data on previous experiences with Willow (in a blended learning context) were mined by a knowledge engineer and analysed by the educators in order to take into account feedback from previous course iterations in the recommendations design.

To provide some insight on the data mining process carried out, we comment on some details on the usage of the Weka [136] implementation of C4.5 decision tree classification algorithm to find out when it seemed to be appropriate to recommend learners to use Willow to review course contents. The idea behind is to deliver the recommendation only when learners are not going to start the review by themselves, to be the less intrusive possible. For this analysis, the class attribute was a Boolean value, which determined whether the learner had used Willow to review the lessons or not. The predicting attributes used were user data (i.e., the learner profile) and usage data (e.g., the number of sessions in the system, visits to course pages and contributions to platform services). The decision tree obtained showed that three sessions seemed to be a relevant value to recommend starting the review of the lesson with Willow as, for the given data from past course interactions, it discriminated between those learners who review concepts with Willow, and those who did not do it.

Thus, collected information report problems that a learner can come across while interacting with Willow and consisted in both qualitative features, such as the fact that some learners did not use Willow to review the lessons, and quantitative data, such as learners accessing at least three sessions had used Willow to review the lessons. With this information, and following the scenario-based approach [137], next in the activity requirements specification a problem scenario was produced. Recommendations were identified in that problem scenario to turn it into the so called solution scenario. For instance, in our example, recommendation R2 (whose final description is detailed below in Table 2) was identified aimed to achieve that all learners use Willow to review the lesson. Each of the recommendations suggests the learner to carry out a specific action on a given learning service of the system when certain conditions (which can refer to the learner profile, situation in the course or interaction with the platform services) occur. These recommendations aim to tackle a specific educational rationale, very often related to awakening meta-cognitive issues during the learning process.
After that, in the activity create design solutions, a focus group with four additional educators was established to review the recommendations obtained in the solution scenario, classify them with a card sorting task, discuss their educational value and unify the recommendations wording. Regarding R2, participants agreed on the relevance of this recommendation and considered plausible the number of sessions proposed as applicability condition. For completeness of the recommendation logic, they added two additional applicability conditions to be consistent with the nature of the recommendation: 1) the learner has to be enrolled in the course, and 2) the learner has not accessed the reviewing functionality yet.

Finally, in the activity evaluate design against requirements the recommendations proposed by the educators and refined in the focus groups, were validated with a Wizard of Oz with potential users (15 learners and 15 educators). They were shown how recommendations would look like when delivered in Willow and asked to evaluate their perception and attitudes towards them. Here, R2 was classified in the category ‘active participation’ and received an over-average rating both by the learners and educators.

As a result, 12 recommendations were designed. They were semantically characterised in an incremental way along the design process, following a similar approach than the one used to elicit recommendations for a course for getting started in using dotLRN LMS [7]. Table 2 collects the recommendations’ type (i.e. ‘what’ should be recommended), their applicability conditions (i.e. to ‘whom’ and ‘when’ the recommendation should be offered) and their educational rationale (i.e. ‘why’ the recommendation is to be delivered). Regarding the example provided for R2 design, it has to be mentioned that educators used the outcome from the data mining process as the initial value for the applicability condition, being aware that it could be mere illustrative since it was obtained from data gathered in a blended learning scenario, which might differ from the current e-learning one. Thus, the applicability conditions might need to be revised with the outcomes after the recommendations delivery in the e-learning context. This revision is considered as part of the TORMES methodology, as commented below. In fact, as commented later in section 4.5, changes to R2 applicability conditions were proposed in the auditing phase from the analysis of the interaction data. This outcome does not limit the value of the UCD approach proposed. On the contrary, it empowers its potential to design valuable recommendation in an iterative way involving users (both educators and learners) along the process.

<table>
<thead>
<tr>
<th>ID</th>
<th>Rec. type</th>
<th>Applicability conditions</th>
<th>Educational rationale</th>
</tr>
</thead>
<tbody>
<tr>
<td>R1</td>
<td>Choose a lesson to review</td>
<td>Learner is in the course. Learner has not visited the page where lessons can be selected.</td>
<td>Help the learners to focus on their study.</td>
</tr>
<tr>
<td>R2</td>
<td>Start the review</td>
<td>Learner is in the course. Learner has not visited the page where the questions can be answered (reviewing concepts functionality). Learner has spent less than three sessions in the course.</td>
<td>Learners review the contents of the lesson by answering the questions formulated by Willow.</td>
</tr>
<tr>
<td>ID</td>
<td>Rec. type</td>
<td>Applicability conditions</td>
<td>Educational rationale</td>
</tr>
<tr>
<td>-----</td>
<td>---------------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>R3</td>
<td>Review the concept estimated as less known</td>
<td>Learner is in the course. Learner has answered at least one question. There is a concept in which the learner has a confidence value below 0.20. Learner has not visited the page with the definition of that concept.</td>
<td>Focus the study on less known concepts, instead of reviewing again other concepts better known to be more effective in the learning process.</td>
</tr>
<tr>
<td>R4</td>
<td>Use the forum to share a doubt</td>
<td>Learner is in the course. Learner has answered at least two questions. Learner has not contributed in the forum.</td>
<td>Foster the communication with other learners to share doubts and get or offer help.</td>
</tr>
<tr>
<td>R5</td>
<td>Read a thread of the forum with many posts</td>
<td>Learner is in the course. There is a forum thread with at least ten messages. Learner has not visited the thread.</td>
<td>To motivate learners to read these long threads as they may have important information for the course. By reading these long threads, she may also be motivated to post.</td>
</tr>
<tr>
<td>R6</td>
<td>Read the educators’ welcome message</td>
<td>Learner is in the course. Learner has not visited the welcome message posted by the educator.</td>
<td>To read the instructions provided by the educators to follow the course. It emulates the first face-to-face introductory class of any course in which the educators provide the instructions of the course.</td>
</tr>
<tr>
<td>R7</td>
<td>Change the avatar that represents Willow</td>
<td>Learner is in the course. Learner has not visited the page where the system avatar can be changed.</td>
<td>Make the learner aware of the philosophy of the platform based on the dialogue metaphor between Willow and the learner. It also strengthens the learner meta-cognitive capabilities in the learning process.</td>
</tr>
<tr>
<td>R8</td>
<td>Change the avatar that represents the learner</td>
<td>Learner is in the course. Learner has not visited the page where the user avatar can be changed.</td>
<td>Make the learner aware of the philosophy of the platform based on the dialogue metaphor between Willow and the learner. It also strengthens the learner meta-cognitive capabilities in the learning process.</td>
</tr>
<tr>
<td>R9</td>
<td>Look at the learner conceptual model</td>
<td>Learner is in the course. Learner has spent at least one session in the course. Learner has already answered four questions. Learner has not visited the page where the personal conceptual model can be visited.</td>
<td>Motivate the learner to keep answering more questions so that more concepts are marked as known. It also strengthens the learner meta-cognitive capabilities in the learning process.</td>
</tr>
<tr>
<td>R10</td>
<td>Look at the conceptual model of the class</td>
<td>Learner is in the course. Learner has spent at least two sessions in the course. Learner has not visited the page where the class conceptual model can be visited.</td>
<td>Motivate the learner to keep reviewing if her conceptual model has more unknown concepts than the class conceptual model or, to congratulate learners who have more concepts estimated as known as they are excellent learners (above the average of the class) so that they keep being the top of the class. It also strengthens the learner meta-cognitive capabilities in the learning process.</td>
</tr>
<tr>
<td>R11</td>
<td>Log in Willow to start the course for the first time</td>
<td>Learner is enrolled in the course. Learner has not entered in the platform. The recommendation has not been given in the current day.</td>
<td>Recall the learner that the course is activated, and that there is a deadline in which the course will finish so she should log in Willow to study as soon as possible.</td>
</tr>
</tbody>
</table>
Table 2. Recommendations elicited for the course EBIFE

Recommendations elicited and reported in Table 2 take into account education issues. Each of them is semantically modelled and ready to be published in the SERS that provides the recommendation service to the LMS.

As a result of this phase, both the packaged course and the 12 recommendations defined in terms of if-then rules were ready to be instantiated.

4.3 Publication phase

In this phase, a learning space for the course was created in Willow. Packaged contents were imported and the 12 recommendations elicited and modelled in the previous phase were instantiated. The recommendations designed in Table 2 were implemented as the rules defined in Figure 3 in terms of if-then sentences where the applicability conditions define the selection mechanism. Learning services were properly configured. In particular, the welcome message and three forum threads were created. Previous to enrolling learners and educators, a usability and accessibility expert evaluated the environment (after the aforementioned preparation by the educational support officer).

<table>
<thead>
<tr>
<th>ID</th>
<th>Rec. type</th>
<th>Applicability conditions</th>
<th>Educational rationale</th>
</tr>
</thead>
<tbody>
<tr>
<td>R12</td>
<td>Log in Willow to keep reviewing</td>
<td>Learner is enrolled in the course. Learner has not entered in the course in the last two days. The recommendation has not been given in the last two days.</td>
<td>Recall the learner that the course is still ongoing, and that there is a deadline in which the course will finish so she should log in Willow to study as soon as possible.</td>
</tr>
</tbody>
</table>

Table 2. Recommendations elicited for the course EBIFE

Regarding recommendations publication, at the stage of this experience the tool proposed in [10] was not integrated in Willow. Thus, this process had to be done manually through the database manager of the persistence layer.
In the end, the recommendations designed in the previous phase were published and ready to be delivered when learners log in the learning space to work with the contents prepared and use Willow’s learning services.

4.4 Use phase
This phase deals with the learners’ interactions with contents and services in the e-learning system, where recommendations can be offered to learners whose runtime context matches recommendations applicability conditions as defined by the rules presented in Figure 3. In order to evaluate the support provided, a usability and accessibility expert tracked learners’ behaviour with empirical testing methods to learn about their interactions. Both quantitative data (from data logs) and qualitative data (from pre and post test, and satisfaction questionnaires) were gathered.

In the first edition of the EBIFE MOOC, a total of 182 participants were enrolled, but 9 of them left the course before the end for personal issues (health problems or too much work at their job). Regarding their profile, most participants were in the age range 35-55 years (79%) and there were more women than men (45% male vs. 55% female). One of the participants reported having visual impairments.

To compare the benefits of receiving recommendations, previous to enrolling in the course, participants were randomly divided into two groups. Half of them were assigned to the experimental group (EG), and thus, offered recommendations (if applicability conditions match learners’ context), and the other half were assigned to the control group (CG), and thus, were not offered recommendations at all. Before the second module took place, learners were again randomly grouped into the EG and CG, having in the end four different groups: EG-EG, EG-CG, CG-EG and CG-CG.

In order to evaluate recommendations impact on learners, system logs with learners’ interactions were observed to check if the recommendations were properly delivered. Learners in EG followed 348 recommendations in M1 and 166 in M2. A statistically significant analysis was carried out on learning performance measures to assess if there was a significant impact (with confidence level set to 95%) when comparing the EG and the CG in each module. The Fisher's test was used for categorical data and the unpaired t test for continuous data. Indicators considered to evaluate the learner performance are compiled in Table 3, showing their goal and the measures to compute them.

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Goal</th>
<th>Measures</th>
</tr>
</thead>
<tbody>
<tr>
<td>Learning effectiveness</td>
<td>Measures the achievement of the learning goal.</td>
<td>- How intensively the system has been used for answering questions: average questions answered (continuous data) and answered all questions (categorical data)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Participants have done the post-test, which is not mandatory: post-test done (categorical data)</td>
</tr>
<tr>
<td>Learning efficiency</td>
<td>Measures the resources used to reach the learning goal until the module is finished and number of activities that are successfully completed in time.</td>
<td>- Number of sessions needed to get to the end page of the module: sessions to get to the end page (continuous data)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Post-test filled in during the length of the module: post-test in time (categorical data)</td>
</tr>
<tr>
<td>Engagement</td>
<td>Measures the involvement of learners in the module in terms of their connection behaviour.</td>
<td>- Participants who connected everyday: had sessions every day (categorical data)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Time required to enter the module for the first time: days to enter the module (continuous data)</td>
</tr>
<tr>
<td>Knowledge acquisition</td>
<td>Measures if learners have improved their knowledge after taking the module.</td>
<td>- Knowledge gain, comparing the outcomes from the pre-test and the post-test</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Results from the module activities: average of correct answers (continuous data) and all answers correct</td>
</tr>
</tbody>
</table>
A positive statistical impact was found on indicators dealing with the *engagement*, specifically more participants from those who received recommendations had sessions every day of the course in M1 (categorical data: sessions every day in EG=30.59% and CG=18.91%, \( p=0.0299 \)) and participants who received recommendations waited less days in average to enter in M1 (continuous data: average days in EG mean=1.36, SD=3.86 and CG mean=2.19, SD=6.31; \( p=0.0206 \)). Recommendations have also impacted on the *learning efficiency*, specifically participants who received recommendations required in average less sessions to get to the end page in M2 (continuous data: average sessions in EG mean=4.21, SD=6.26 and CG mean=5.68, SD=6.39; \( p=0.0462 \)). Regarding knowledge acquisition and learning effectiveness, the global analysis (i.e. CG vs. EG outcomes) did not show a statistical difference in any of the indicators computed. However, when comparing participants being recommended a certain recommendation versus participants not being recommended that recommendation, some statistically significant effect was found (with \( p<0.05 \)) for the recommendations (except for R11). In particular, participants receiving a recommendation answered correctly more questions (*knowledge acquisition*), as well as they also answered more questions and more of them did the post-test (*learning effectiveness*). This result suggests that a global significant impact might be caused on the knowledge acquisition and the learning effectiveness, since the only recommendation that do not report any significant impact (i.e. R11) does not focus on the learning process itself, but on entering in the module.

For those participants in the EG, the feedback voluntarily provided through the dialogue with the system (as previously shown in Figure 2) was also analysed. Almost 40% (127/348) of the recommendations were fed back in M1 and 45% (75/166) in M2. Over 80% of recommendations followed were found useful (84% in M1 and 82% in M2) by participants. When analysed individually, there were three recommendations which were not as well rated as the others: R4 (75% in M1 and 67% in M2), R7 (63% in M1 and 67% in M2) and R8 (64% in M1 and 33% in M2). Top rated recommendations were R9 (100% in M1 and M2) and R10 (91% in M1 and 92% in M2).

Moreover, all learners, both in EG and CG, were asked at the end of the MOOC, about their perception of having been recommended, especially if changes were perceived by those who received recommendations just in one of the modules. Results showed that participants did not differentiate the feedback provided by Willow from the recommendations functionality added, as it was desired in order to comply with the usability consistency principle. Since recommendations were not detected as an external functionality, they seem to have been introduced in Willow in an unobtrusive way.

The system usability was computed with the standardised SUS questionnaire [138]. The average scores are high in all groups, as follows: 1) EG-EG: mean=74.40; SD=15.29; 2) CG-EG: mean=75.09; SD=13.68; 3) EG-CG: mean=71.21; SD=17.55; 4) CG-CG: mean=73.61; SD=16.73. The ANOVA analysis do not report a statistical significance for alpha=0.05 (\( p=0.8068, F=0.32 < F_{crit}=2.69 \)). An average rating of 68 has been abstracted from over 500 studies [139], which means that a SUS score above a 68 (like in this case) is considered above average. This confirms the results of the usability evaluation carried out in the publication phase, where no problems were found when a usability expert reviewed the environment settings (Willow plus contents and
recommendations). In addition, the participant with visual impairment finished the course without reporting any accessibility issues.

4.5 Auditing phase
The outcomes gathered in the use phase were analysed by educators and adjustments to some recommendations, either in text, applicability conditions or both were proposed for the next iteration design phase. This second iteration (EBIFE 2) involved 204 participants. The justification for the proposed changes to some modifications is explained next, together with the corresponding outcomes of the second edition of the MOOC, where the revised recommendations were delivered. Settings both in EBIFE 1 and EBIFE 2 were the same, except for the participants (new and different users were enrolled) and the recommendations redesigned commented here.

The applicability conditions in R2 were changed, as when analysed individually, R2 did not show any statistical effect in learning effectiveness, efficiency and knowledge acquisition impact except for course engagement. This issue was previously considered as a possible finding by the educators, as the value for the applicability condition was obtained from the analysis in the blended learning context. Thus, educators decided to offer the recommendation only when learners had already answered some questions independently of the number of sessions spent, to make learners review recommendations any time, not only in the first sessions (as in the design done for EBIFE 1). They also improved the textual description of the recommendation following the changes made in the applicability conditions. Changes made were effective, as a significant statistical positive effect was registered in EBIFE 2 M1 for effectiveness (answered all questions: EG=34.38% and CG=12.41%, p=0.0063; average questions answered: EG mean=53.61%, SD=15.16% and CG mean=24.63%, SD=10.90%, p=0.0000; post test done, including the extra days give after closing the course: EG=84.38% and CG=48.97%, p=0.0003), efficiency (post test done before the closing of the course: EG=75.00% and CG=31.03%, p=0.0000) and knowledge acquisition (all correct answers: EG=34.38% and CG=17.72%, p=0.0054; average correct answers: EG mean=49.10%, SD=17.64% and CG mean=18.36%, SD=11.99%, p=0.0000).

Applicability conditions in R3 and R7 were also revised because of the difference in the context considered (blended learning vs. e-learning). In particular, after EBIFE 1, educators decided to reduce to 0.10 the value for the confidence level to trigger R3, and decided to limit the delivery of R7 after the learner had spent 4 sessions in the course. This arbitrary number of sessions reflects educators’ thought after EBIFE 1 experience on the better delay for that recommendation.

Finally, some changes were also done to the description of R4, R7 and R8, and these turned out as good decisions, since the ratio of useful recommendations increased in EBIFE 2 for those recommendations (from less than 75% to over 86%).

Thus, the analysis carried out after the first edition of EBIFE MOOC has been valuable to identify recommendations whose design could be (and was) improved in the next round. This shows the benefits of involving the educators along the eLLC to guarantee that the adaptation produced (in this case the recommendations) are designed taking into account the learners’ needs, and are supported with appropriate UCD methods. The recommendations design is iteratively improved. When there is no data from previous similar course runs, related courses can be considered as input for eliciting
recommendations opportunities, although the applicability conditions might be revised after the course execution in the new context.

5. Discussion

From the outcomes reported in this paper follows that UCD methods applied along the eLLC have shown their value for enriching Willow with educational-oriented recommendations. These recommendations have provided appropriate navigational and educational support in a MOOC scenario (involving nearly 400 learners). The approach is grounded on extending the adaptation capabilities of this web-based educational system with adaptive navigation support by highlighting potential educational based useful actions in that system to offer in an e-learning context educational-oriented recommendations that have improved the learning experience. The approach includes a UCD methodology that covers the eLLC as the recommendations have been modelled, instantiated in the system, used by the learners and analysed the feedback obtained from their usage in order to assess their validity.

Recommendations designed were delivered following a service-oriented approach aimed to support interoperability and reusability in e-learning. This requires the semantic modelling of the recommendations, which were described in terms of ‘what’ should be recommended (object and action), to ‘whom’ and ‘when’ the recommendation should be offered (applicability conditions) and ‘why’ the recommendation is delivered (rationale). It also includes specific recommendation features (i.e. recommendations origin, a categorisation, the relevance and the course stage).

We have shown how user centred concerns arose and were confronted while extending Willow web-based educational system with adaptive navigation support in terms of recommendations. Following the eLLC, educators were supported with UCD methods (e.g. interviews, user observations, scenarios of use, card sorting, focus groups, Wizard of Oz) to design recommendations that cater for the learners’ needs in an unobtrusive manner and awaken meta-cognitive issues during the learning process. Benefits come out from recommendations utility, which resulted in an improvement on the learning performance. Most of the participants perceived recommendations as useful, and they were not perceived as an external functionality to the system, while keeping the high usability level of the system without recommendations. Thereby, the process has followed the usability design principle of consistency, which guarantees users’ unawareness on the addition of an external system functionality (i.e. recommendations). Otherwise, users’ interaction with the system would have experienced an added complexity that would have affected their natural interaction flow.

As a result, educators were able to translate their instructional designs best practices to the adaptation process and get feedback from the application in a real world scenario. Recommendations usage was evaluated by the educators, who accordingly redesigned some of the recommendations for the next iteration of EBIFE MOOC. The adjustments proposed showed their benefits. In this way, it can be seen how personalisation can be supported in an iterative way along the eLLC, where recommendations provide dynamic adaptations in the use phase (runtime) in terms of adaptive navigation support.

In order to provide practical guidelines to other researchers for extending their LMS with adaptive navigation support, we have compiled in Table 1 the key aspects of our UCD approach. For each of the eLLC phases, Table 1 presents the phase goal, the tasks
to be carried out, the actors involved, the input to be taken into account and the output produced. Several actors are identified: learner, educator, usability and accessibility expert, educational support officer, knowledge engineer and software developer. Usability and accessibility experts are required to carry out heuristic and accessibility evaluations on the learning infrastructure (system, contents, space) as well as to perform the empirical testing to track learners’ behaviour in the use phase. Educational support officers can help with the course packaging and e-learning system administration issues. Knowledge engineers can extract implicit knowledge from past course interactions. Educators are expected (besides preparing the learning design of the course) to take an active role in both the recommendations elicitation process and the analysis of the recommendations impact on the learning experience. Usability methods such as heuristic evaluations, gathering and elicitation methods and empirical testing are considered in the proposed UCD methodological approach. Thus, this paper presents a contribution to the current urge of building personalised e-learning systems, which consists in reporting where UCD methods can be applied along the eLLC to design and evaluate personalised support through recommendations in web-based educational systems.

In a MOOC context, where the large number of learners and course instantiations require significant tutoring resources, recommendations can provide timely personalised support that reduces the educators’ workload involved in assisting the learners’ throughout the course. However, a trade-off exists, as educators should design in advance this support. Nevertheless, in our view this trade-off is reduced if both UCD methodologies and standards and specifications are used. The former provide tools to optimise and guide the educators’ involvement along the eLLC phases while the later facilitate reusability of contents not only from instantiations of the same course, but even when the e-learning system that support the course changes. Further, it also considers data mining techniques that can be used to enrich the knowledge obtained through UCD methods. Thus, the SERS approach presented here would expand their potential when e-learning systems are fully developed in terms of standard-based service-oriented architectures, which facilitates that recommendation services and data mining analysis provided by external components can be easily integrated.

Furthermore, technological infrastructures to support educators along the eLLC to manage the recommendations elicitation, design, publication and evaluation of their effect on the learning are demanded in order to move this approach from research to daily educational practice. Investment is currently done worldwide in developing platforms for MOOC and although it is debatable which is the best approach, figures show that there are thousand or even hundreds of thousands learners involved [140] that require educational-oriented personalised solutions.

6. Conclusions and Future work

This paper has presented a user centred design approach for enriching web-based educational systems with educational-oriented recommendations. The approach is built from aDeNu experience in past research projects (i.e. aLFanet, ALPE, EU4ALL). The key finding from them was that building personalised learning e-environments is a process that has to consider the learners’ needs during the whole e-learning life cycle. To cope with this need, in this paper we have reported the usage of user centred design methods along the e-learning life cycle to designing and evaluating personalisation support through recommendations in web-based educational systems. Moreover, in
Table 1 we have compiled the key aspects of our user centred approach in a way that can be used by other researchers when designing personalised e-learning systems.

The proposed approach has been tested in a real world e-learning scenario through an adaptive system called Willow that has run a Massive Open Online Course (MOOC) with nearly 400 learners (173 in the first edition and 204 in the second). With the help of TORMES user centred design methodology, two educators have taken an active role in the recommendations elicitation process during the design phase. They have also been supported by a software developer, a usability and accessibility expert, a knowledge engineer and an educational support officer along the whole e-learning life cycle, as outlined in Table 1.

The main lesson learnt from this experience is that there are apparent advantages derived from following a user centred design approach along the e-learning life cycle. Through user centred designed recommendations given to learners, the e-learning system addressed educational issues, such as i) overcoming the limitations of the course design and/or the platform features, or ii) pointing to relevant information from the large amount of contributions provided by learners and educators.

On going works at aDeNu are focused on considering affective issues in the adaptation support along the e-learning life cycle since the inclusion of this kind of issues is expected to increase motivation and benefit the achievement of learning outcomes [141]. This work is being carried out in the context of the MAMIPEC project (TIN2011-29221-C03-01) [142], where recommendations are being used to provide adaptive inclusive personalised support in educational scenarios [143] that deal with intelligent contexts which consider multimodal approaches for affective modelling [144]. Moreover, affective issues in collaborative learning scenarios are also being addressed [145].

Regarding recommendations applicability in MOOCs, we also intend to take advantage of this user centred design approach to identify and evaluate valuable domain independent educational-oriented recommendations that can be generally applicable in MOOCs. Moreover, although there exist standards and specifications to describe the educational elements involved in the recommendations process, and developing efforts are done towards a modular approach in the next generation of learning management systems that are to support service-oriented architectural approaches, there is still a need to further research on the software engineering aspects required to fully support semantic educational recommender systems. It is expected that the growing interest in delivering MOOCs would facilitate that support.

Acknowledgements

The research outcomes presented in this paper have been possible thanks to the funding obtained in several projects. For Willow, “Ubiquitous Collaborative Adaptive Training” (U-CAT, TIN2004-03140). For the SERS approach, “Active Learning For Adaptive internet” – aLFanet (IST-2001-33288), “Accessible e-Learning Platform for Europe - ALPE (eTEN -029328)”, “European Unified Approach for Accessible Lifelong learning” - EU4ALL (IST-2006-034478) and “Multimodal approaches for Affective Modelling in Inclusive Personalised Educational scenarios in intelligent Contexts” - MAMIPEC (TIN2011-29221-C03-01).
References


