Doctoral Consortium at ICWL 2016

Maria De Marsico[1], Zuzana Kubincova[2], Carla Limongelli[3]

[1]Sapienza University of Rome, Rome, Italia
demarsico@di.uniroma1.it
[2]Comenius University in Bratislava, Slovakia
kubincova@fmph.uniba.sk
[3]University of Rome Three, Rome, Italia
limongel@dia.uniroma3.it

1 Introduction

Web-based learning has revolutionized distance education from the foundations. The tremendous variety of new possibilities offered has also caused the upgrade of methodologies and strategies entailed in face-to-face education. The possibility to exploit remote and distributed resources, multimediality and hypermediality, all allow establishing unprecedented dynamic links among traditional chunks of knowledge and novel ways of representing information. The use of social networks further spurs discussion, exchange of ideas and experience, and cultural mix-up. This supports the concrete implementation of socio-constructivism in a way that was unexpected by its pioneers. However, every medal has a reverse side. Relying uniquely on remote deployment of knowledge and configuration of learning tasks suffers from the lack of personal, concrete interaction both among learning peers and with teachers. It is often the case that distance learning becomes a lonely experience, with no factor actually pushing towards communication. Looking at educational forums and blogs, though crowded with messages and threads, it is possible to notice the low number of active participants with respect to the enrolled ones. Learning material is serialized according to standard criteria, and the higher form of adaptation that is often implemented is the one represented by barriers among a learning node and the next one that are regulated by the individual achievements (typically, a sufficient mark achieved in a multiple-choice test). Special needs are addressed in the most immediate way, often without taking an additional effort to really understand difficulties on a per student basis. In traditional classroom education, most such problems were addressed, when applicable

of course, by the experience and sensibility of teachers, that are able in some cases also to anticipate learning problems and to take personalized solutions. Of course, this is true only according to the teacher's qualities, but in distance learning there is hardly a possibility to carry out similar strategies.

The papers presented for this Doctoral Consortium tackle two specific and very interesting problems by addressing different relevant aspects. The first one is related to the way to exploit the interesting possibilities offered by web based learning, namely the retrieval of contents integrating the teachers' lessons [1][3]. However, finding such content might be a non-trivial task. Especially designed expert systems and agents can help. Such systems can help both students in deepening their study, and teachers in developing their courses. The second group of papers deals with the use of gamification and game-based learning to enhance the involvement of learners into educational activities [2]. Both problems entail the use of state-of-the-art technologies and methods. From one side, web based learning offers possibilities of paramount value, but from the other side some problems still need to be addressed and solved. In this context, researchers, and in particular PhD students, represent the front line of investigations aiming at improving the quality and results of this revolutionary learning approach.

The papers have been reviewed from two anonymous referees and all the manuscripts have been revised. This preface summarizes both paper contents and the most relevant comments raised during the discussion at the DC meeting.

The first group of three works includes the paper by Carlo De Medio with title "An Intelligent Agent with Ontological Knowledge: Classification of Educational Materials to Support the Creation of Online Courses". It aims at providing teachers with an intelligent assistant to help courses configuration. The study starts from determining the dependencies between topics related to the learning objects in a repository (LOs) by inspecting related resources in Wikipedia, in order to understand the prerequisite and successor dependencies between materials. The author envisages two modalities to exploit the planned system. In the assisted modality, the teacher inserts an initial concept and receives back a list of possible next LOs; a selection in this list corresponds to the generation of a new list of concepts related to the one selected until the completion of the course. In the automatic modality, the teacher inserts an initial concept and the intelligent assistant automatically creates a course according to teacher styles. A relation Classifier is proposed that is able to determine the precedence between a pair of LOs, according to a number of features that relate to both single LOs and to the pair.

During the discussion, two main issues were raised.

The first concerns the significant dependence from Wikipedia as a unique reference: how can a single concept be used to automatically create a course? As a consequence, another basic issue arises, that concerns the definition of the term "concept". How "large" a concept is? Is a concept as large as a general topic or as small as more detailed arguments? This issue cannot find a simple answer, this research is still at the

beginning and a careful experimentation will allow adjusting the definition of the term "concept" that better fits for this purpose. The final objective of this PhD is to integrate this mechanism for finding dependencies in a Learning Management System and tested by a community of Teachers.

The second remark concerns the reason why length of text in LOs (if any) plays such a relevant role in the feature set used for precedence identification. It is interesting to note that this consideration comes from the observation of the experimental results. They have confirmed that the length of a LO, in terms of a number of words, is one of the most relevant features to determine dependencies between two Learning Objects.

The work of Matteo Lombardi is titled "Discovering Educational Resources on the Web for Technology Enhanced Learning Applications" and is related to automatic discovery of relevant educational resources from the Web. In practice, the starting point is to explore the possibility to use the entire Web as a repository of teaching material. However, the retrieval of resources from this enormous and unorganized space presents some critical issues. The main ones are summarized by the research questions implied by this work: which features can reliably describe an educational resource, how resources can be extracted from heterogeneous sources, and how they can be represented to allow fast and easy access. The option of relying on popular standards for the description of Learning Objects (LOs) is often hindered by issues such as the diversity of metadata standards and their lack in representing some relevant educational aspects. One of the aims of the PhD work is to propose an automatic extraction of educational features, able to represent two groups of data: the characteristics proper of the resources, and the context where the resource has been used (including other related resources). This should feed an appropriate classification strategy, able to drive web crawling and to overcome present limitations of current crawling approaches. As a matter of fact, these approaches are tailored on topics, terms and domains, but not on the context of usage of Web resources, i.e., teaching/learning. To address this limitation, the PhD project proposes the development of a suitable classifier to be integrated into the crawling process. The classifier is developed by taking as a reliable model the resources hosted on trustworthy educational platforms, and therefore already used in an educational context. During crawling, it should be able to compare the new resources coming from the Web with the models and decide if they are appropriate for education. Discovered knowledge is represented as a concept map, where resources are entities connected by educational relationships. The aim is to offer better educational applications both to students and to educators.

Both the reviews for this paper and the following discussion during the DC meeting pointed out that, even though the underlying motivations of the work presented in the paper have been investigated for years, its originality relies on the methodologies to be developed in order to identify and crawl educational resources on the Web. The use of a classifier in connection with the crawling process was especially appreciated. However, some difficulties were also pointed out. The first one is related to the target of the approach, i.e. resources structured according to available metadata standards. Actually, the web is full of unstructured and yet extremely useful data, that at the moment

is left out from the process. Even if nowadays the Web is not only composed by unstructured information, and structured and semantically annotated web pages begin to appear at a high pace, the great majority of multimedia educational data is still available "in the wild". A further issue raised during the meeting regards the licensing of Web resources, since not all assets available on the Web can be freely re-used in educational settings. In practice, the performance of the proposed strategy, in terms of the actual amount of collected resources, might be below the expected and achievable one if these problems, especially the former one, are not taken into account. However, processing of such huge amounts of unstructured information for assessing educational suitability is still a perhaps too ambitious goal given the present algorithms and computational resources.

The work of Alessandro Marani with title "WebEduRank: an Educational Ranking Principle of Web Resources for Teaching" won the ICWL 2016 Doctoral Consortium Best Proposal Award and is part of the same overall project as the previous one. It addresses the recommendation of web resources for teaching. The solution proposed aims at retrieving resources not only from dedicated Learning Object Repositories (LORs), but also from MOOCs, as well as from any other web sources. The project deals with the design of WebEduRank, a ranking principle for rating web resources according to the teaching requirements and needs. The first step of the research has been to identify the attributes that are expected to describe resources. Using these attributes, the design of the WebEduRank has been completed and the methodology for its validation is already devised and implemented. Since the recommendation system is especially designed for teachers, their teaching requirements and experience make up the teaching context exploited to propose a ranking principle of web resources. In particular, the teacher's knowledge involved in the process of teaching (called *teaching knowledge*), includes categories such as Content Knowledge, Pedagogical Knowledge and Pedagogical Content Knowledge. Further attributes can be included in the Instructor Profile. An initial definition of the Teaching Context (TC) of a course includes Prerequisite Knowledge (PK), Concept Name (CN), Course Title (CT), Education Level (ED-LEV), Difficulty (DIFF), Starting Knowledge (SK), Target Knowledge (TK). After this identification step, the research question to answer regards which of these attributes is considered actually relevant and useful by teachers when looking for a resource to include in one's course. To this aim, a questionnaire was submitted to 33 teachers and its outcomes were analyzed. According to such analysis, during the search of teaching resources on the web, a teacher starts from a concept (CN) for a particular course (CT), when some prerequisite knowledge (PK) is assumed to be already learnt by the students. Of course, the teacher will look for resources that comply with the education level (ED-LEV) of the class and with a certain level of difficulty (DIFF). The course may also be a bit different in terms of concepts taught in it because of the starting knowledge (SK) of a class. Finally, the instructor knows, what is the target knowledge (TK) of the course. The next challenge regards the identification of the parts of a web page where to look for that information, including the title, the body, links and headers/highlights. The work proposes to exploit the Expectancy Appearance Matrix (EAM) for representing the expectation that an attrib-

ute appears into a component of the web page. The rows of the matrix are the four components of the web pages analyzed by the WebEduRank, and the columns are the attributes of the Teaching Context. Each matrix element is the likelihood that the corresponding attribute is found in the corresponding part of the web page. The events of finding or not an attribute in a certain part of the web page are independent of each other. Tuning of the system to discover the most appropriate values for each element of the matrix is carried out by the analysis of the dataset DAJEE of resources extracted for Coursera platform. The paper then presents a thorough assessment of the suitability of the proposed approach. The completeness of the analysis and the derived system validation are further indicators of the maturity and robustness of this work

During the presentation it was underlined that the proposed principle is not much focused on the ranking of the resources according to a topic. As a matter of fact, Google can do this very well and with a remarkable performance. The main problem addressed by WebEduRank is the ordering of a set of web resources according to their suitability for teaching in a certain context, and according to the characteristics of a specific teacher. In other words, this is the symmetric task with respect to personalized learning from the point of view of learners. Due to the dependency on keywords and topics, Google is not able to act as a recommendation system with these features. However, the presented system has a limitation too that was pointed out. One of the relevant parts of a web page used to extract information is the document body. However, in a multimedia resource this part could not be a textual one (e.g., it might be a video) and no links/highlights might be present. This would reduce the analysis to the title only, and this may cause a wrong ranking. A possible solution might be to take into account the set of "neighboring" pages that belong to the same information structure as the analyzed one. According to the author, this problem will be tackled by future work on the system.

The second group of proposals includes the work by Mehdi Rizvi with title "Children's Socio-Emotional Learning Enhanced by Tangibles for Group Activities".

This PhD research focuses on interdisciplinary topic of socio-emotional learning (SEL) of children. Its main goal is to design and evaluate tangible interactive objects to enhance children's social-emotional learning experience. The PhD student is a member of an interdisciplinary team in which experts from the fields of interaction design and education are represented.

To accomplish the research goal two research questions are defined. The first one addresses the problem of how to design the tangibles that enhance children's experience with SEL group activities, focusing on conversation. This problem is solved by the use analysis and the design of prototype tangibles that are assessed in sessions with the children. The User eXperience design together with the lean approach are used as the methodology for tangibles development.

The second research question deals with the same problem as the first one however, from the perspective of enhancing the educators' experience. To deal with this problem and be able to consider their needs, the educators are involved in the tangibles design process, in their analysis and evaluation.

The paper introduces one design case of a tangible: TurnTalk which was designed to engage children in a balanced conversation as well as to promote the children's awareness of the turn-taking rule and reflection on conversation patterns.

Although this project is at an early stage several results were published in two conference papers already. This PhD research seems to be original and promising.

The reviewers agreed in considering the promising idea to develop a tangible for conversation of children in the turn-taking conversation paradigm. In addition, there is few research in literature regarding the second research question. It is interesting to notice that, though the main stakeholders in educational processes are both the teacher and the learner, the former are often left out from the design of learning tools, and often their opinions and preferences are simply neglected. In general, one question that is often raised during discussions and presentations of new teaching/learning tools is, whether the teachers deem them really useful. In many cases this is not clear, and the new proposal is a mere technical exercise. In this case, teachers were actually involved and their response was surely positive.

The obvious observation raised during the presentation is that, in order to take advantage of the devised strategies in web-based learning, the concept of "tangible" must be translated into digital/virtual terms to be used at a distance, e.g., on the web. Nowadays, this is possible through two different strategies. The first one entails devising an equivalent interaction paradigm exploiting traditional input/output devices, yet providing a similar interaction pattern. This might be easier from a purely technological point of view, since no special equipment is required. However, due to the intended goal and to the specific characteristics of target users (children), it might result quite tricky to obtain similar results maintaining the same level of interest, attention, involvement and also enjoyment. The second strategy entails exploiting the potential of virtual reality and natural/tangible interfaces. In this case, mapping onto the distance setting may result more natural. It would be not trivial in any case, due to possible difficulties in using equipment, possible latency of communication, and drop of attention due to the lack of physical contact. However, children (digital natives, as defined by Prensky in 2001) are becoming more and more acquainted with communication devices, much more than older users (digital immigrants), and complex equipment is becoming more widespread, usable and cheap. It will be interesting therefore to follow the future developments of this work.

The second work in this group is the proposal of Annalisa Terracina with title *"*Game@School. Teaching through Gaming and Mobile-based Tutoring Systems".

This PhD research concentrates on the effects of educational games on learning. More specifically, long-term learning is the main target. Notwithstanding the increasing popularity of gamification and game-based learning, there is a lack of empirical evidence of the impact of educational games on learning. Therefore, the core of the proposal relies on the following three research questions: 1) Does Educational Game support learning? 2) Does Educational Game undermine long-term learning? 3) Does the use of Intelligent Pedagogical Agents foster the idea of personalized mentors?

To answer the research questions, the PhD student is designing and developing a serious game for classroom teaching of STEM (Science, Technology, Engineering, and Math) subjects, particularly physics, at high school. The game integrates two advanced items: a 3D virtual learning environment (VLE), and an Intelligent Pedagogical Agent (IPA). The latter is characterized by an emotional intelligence and behaves in different ways according to different emotional contexts. The system offers a shared knowledge forum and an assessment tool for teachers too. The originality of this work resides in integrating all these aspects (Serious Game, VLE, IPA, high school subject, classroom teaching, etc.) into one system. In practice, the research described aims at covering different fields:

- Pedagogy and psychology issues
    - emotion and learning
    - gaming and learning
    - role-play game
    - classroom practice
- Artificial Intelligent and Tutoring System
    - Intelligent Pedagogical Agent
- Virtual Learning Environment
    - 3D Virtual World
    - Teacher's workplace
    - Game engine
- Game Based Learning
    - Educational game

Among the components making up the architecture, the first one (and only at the moment) to be extensively tested has been the IPA. This module is claimed easy to integrate in a stand-alone application, in order to be tested independently from the other game components. The further reported reasons for choosing to test it before the entire game was finished, are absolutely convincing. First of all the IPA is an important component of the game, that can significantly affect the level of students' acceptance. Therefore it was deemed important to have an early feedback on its appreciation before completing the game. As a second point, as beneficial in AI, the algorithm underlying the IPA needed a validation and fine tuning.

This PhD research is interesting but the connection between the design and implementation on the one hand, and the research questions on the other hand should have been better explained. Unfortunately, the work was not presented at the DC meeting, and therefore it was not possible to further discuss these aspects.

To close this preface, we would like to thank the precious work of the reviewers that, besides improving the initial manuscripts, also guided the doctoral students towards achieving their objectives through the best path.

We would also like to thank all the doctoral students for their commitment in the research field of Web-Based Learning.

Finally, a special thank goes to the organizers of ICWL who gave us the opportunity to host our Doctoral Consortium.

1. Atkinson, J., Gonzalez, A., Munoz, M., Astudillo, H.:Web metadata extraction and semantic indexing for learning objects extraction. Applied Intelligence 41(2), 649–664 (2014)
2. Backlund, P., & M. Hendrix, M. "Educational games - Are they worth the effort? A literature survey of the effectiveness of serious games," 2013 5th International Conference on Games and Virtual Worlds for Serious Applications (VS-GAMES), Poole, 2013, pp. 1-8. (2013) doi: 10.1109/VS-GAMES.2013.6624226
3. Romero, C., & Ventura, S. Educational data mining: A survey from 1995 to 2005. Expert systems with applications, 33(1), 135-146 (2007).