The colMOOC project agenda: Towards Integrating Conversational Agents and Learning Analytics in MOOCs

Stavros Demetriadis[1], Anastasios Karakostas[2], Thrasyvoulos Tsiatsos[1], Santi Caballé[3], Yannis Dimitriadis[4], Armin Weinberger[5], Pantelis Papadopoulos[6], George Paligeorgiou[7], Costas Tsimpanis[8], Matthew Hodges[9]

[1] Aristotle University of Thessaloniki, Thessaloniki, Greece
[2] Centre for Research and Technology Hellas, Thessaloniki, Greece
[3] Universitat Oberta de Catalunya, Barcelona, Spain
[4] University of Valladolid, Spain
[5] University of Saarland, Germany
[6] Aarhus University, Denmark
[7] Learnworlds, England
[8] Greek Universities Network, Greece
[9] Telefonica, Spain

Abstract. Higher Education Massive Open Online Courses (MOOCs) introduce a way of transcending formal higher education by realizing technology-enhanced formats of learning and instruction and by granting access to an audience way beyond students enrolled in any one Higher Education Institution. However, although MOOCs have been reported as an efficient and important educational tool, there is a number of issues and problems related to the educational aspect. More specifically, there is an important number of drop outs during a course, little participation, and lack of students' motivation and engagement overall. This may be due to one-size-fits-all instructional approaches and very limited commitment to student-student and teacher-student collaboration. This paper introduces the development agenda of a newly started European project called "colMOOC" that aims to enhance the MOOCs experience by integrating collaborative settings based on Conversational Agents and screening methods based on Learning Analytics, to support both students and teachers during a MOOC course. Conversational pedagogical agents guide and support student dialogue using natural language both in individual and collaborative settings. Integrating this type of conversational agents into MOOCs to trigger peer interaction in discussion groups can considerably increase the engagement and the commitment of online students and, consequently, reduce MOOCs dropout rate. Moreover, Learning Analytics techniques can support teachers' orchestration and students' learning during MOOCs by evaluating students' interaction and participation. The research reported in this paper is currently undertaken within the research project colMOOC funded by the European Commission.

1   Introduction

Understanding and designing with digital technologies has become a relevant competency across disciplines. Computational thinking and programming skills become increasingly important - also for non-programmers. European universities trail behind in offering these transversal competencies for students beyond the domain of computer science. Likewise, the current European Workforce often needs further training for digital literacy. For example, in-service teachers have a need to learn about instructional technologies that were not present during their study times. Last but not least, becoming digital literate may be a crucial springboard for unemployed groups with little access to higher education, e.g. single parents [1].

To help face these challenges, Higher Education Massive Open Online Courses (MOOCs) [2] arose as a way of transcending formal higher education by realizing technology-enhanced formats of learning and instruction and by granting access to an audience way beyond students enrolled in any one Higher Education Institution (HEI). However, the potential for European HEIs to scale up and reach an international audience of diverse backgrounds has not been realized yet. Moreover, although MOOCs have been reported as an efficient and important educational tool, there is a number of issues and problems related to the educational aspect. More specifically, there is an important number of drop outs during a course, little participation, and lack of students' motivation and engagement overall. This may be due to one-size-fits-all instructional approaches and very limited commitment to student-student and teacher-student collaboration.

This paper introduces a new European project called colMOOC that aims to enhance the MOOCs experience by integrating (i) Collaborative settings based on Conversational Agents (CA) both in synchronous and asynchronous collaboration conditions; (ii) Screening methods based on Learning Analytics (LA) to support both students and teachers during a MOOC course and (iii) Reinforce European leadership by forming recommendations and policy guidelines. Conversational pedagogical agents guide and support student dialogue using natural language both in individual and collaborative settings [3][4]. CAs have been produced to meet a wide variety of applications and studies exploring the usage of such agents have led to positive results. Integrating this type of CAs into MOOCs to trigger peer interaction in discussion groups is expected to considerably increase the engagement and the commitment of online students (and, consequently, reduce MOOCs dropout rate). Moreover, this project proposes to use LA techniques as a method to support teachers' orchestration and students' learning during MOOCs by evaluating students' interaction and participation.

A relevant operative goal of the colMOOC project is to develop specific MOOCs for students of the humanities as well as employees and the unemployed to develop digital competencies that have not been covered in basic, mono-disciplinary studies. MOOCs are considered one of the key tools to address the problem of digital illiteracy and equal access to education. MOOCs have shown to not only provide expert and up-to-date knowledge to anyone with internet access, but have also repositioned Universities as beacons of knowledge with impact on society at large [5]. Therefore, the main motivation of the project is to develop MOOCs that would capture the interest of different target groups and develop transversal digital

competencies in specific fields of burning, immediate needs for digital literacy. However, developing MOOCs for lifelong learning that capture the interest of university students, employed and unemployed adults requires innovative, engaging and motivational learning methods. To this end, collaborative learning arrangements and formative feedback can not only improve students' performance, but also minimize chances of dropout and foster engagement. There is an intense need to provide such collaboration settings to enhance active participation and social skills [6]. To achieve this, the project will employ collaborative activities supported by CA.

Considering the above, the main priority tackled in the colMOOC project identifies that there is a need to make MOOCs more collaborative in order to support teachers and engage and motivate students. Moreover, it is critical to ensure that assessment in MOOCs will monitor individual achievement and demand constant engagement [7]. Each learner performs specific individual and collaborative activities during a MOOC course, while his/her performance can be more easily measured and assessed as opposed to a traditional learning setting where individual contributions to a work project are usually unknown or difficult to assess. Furthermore with the usage of learning analytics more meaningful assessment may be achieved, since educators can evaluate the work's progress during each step and adapt the following steps accordingly. This also enables learners to reflect on each specific step and ensure they have obtained the necessary knowledge to progress effectively. Therefore, the project's outcome will allow educators and trainers to apply learning processes where learners participated actively and are engaged throughout the course.

The remainder of the paper is structured as follows: Section 2 reviews the key topics forming the context of the project and in particular the drawbacks and challenges of MOOCs. Section 3 shows the project methodology with the aims and work plan while Section 4 shows the expected results of the project. Finally, Section 5 summarizes the main ideas of the paper and outlines the following steps of the colMOOC project.

2   Background

Supporting MOOCs implies to identify the drawbacks, problems and several user requirements that they have not yet addressed. Since top-ranked academic community joined the MOOC hype, academic sectors have hold controversial discussions on the many MOOC challenges that must be faced before moving on [8][9][10]: (i) high learners´ dropout rate, with only 5% to 15% of participants finishing the courses on average; (ii) limited teaching involvement during delivery stages; (iii) lack of adaptability to a great variety of specific needs.

The colMOOC project aims to provide innovative ways for promoting learners' interaction by enabling the teacher to configure a conversational agent software component which attempts to trigger learners' discussion through appropriate interventions in dialogue-based collaborative activities. The development of the agent will consider all current research evidence on the value of both 'Academic Productive Talk' [18] and the 'Transactive dialogue' [19] frameworks as forms of productive peer dialogue. The main functional features of the agent component include: (a) a user

friendly interface based on the concept map metaphor, for the instructor to easily model the knowledge domain necessary for the agent intervention and also configure the agent behaviour during students' discourse; (b) an appropriately designed MOOC interface for students' synchronous and asynchronous online discussions, where the agent also appears and intervenes in the discussion aiming to trigger productive students' social interaction, domain focused cognitive activity and deeper learning.

Assessment in MOOCs is about supporting student learning and achievement [7]. MOOCs bring a new dimension to assessing such large number of learners. The automated assessments that have evolved in recent years are specifically targeted to assess and evaluated the large enrolments since it is not possible to manually grade and provide feedback to all the learners. Learners can be assessed on time-on-task; learner- course component interaction; and a certification of the specific skills and knowledge gained from a MOOC [11]. While not the primary aim, these assessment techniques will provide an added incentive for the learners to persist and completed the MOOC. Ultimately, the satisfaction gained from completing the course can be potential indicator of good learning experiences. Conversely, enhancing the learning experience can contribute to improving the MOOC completion rates [5][12]. As a result, assessment techniques that permit customization of content catered to the individual learner can track learner behavior and predict learning outcomes. Such a technique will further assist in developing and refining assessment procedures for improving learning outcomes. The colMOOC project will enhance MOOCs capabilities to this direction by exploiting learning analytics based on previous projects and experiences [13].

3 Project methodology

In this section we present the project methodology, including a broad description of the project aims and the work plan, including the evaluation activities.

3.1 Project aims and objectives

The colMOOC project aims to deliver a highly innovative and beyond the current state-of-the-art MOOC model and implementation with the integration of services based on conversational agents and learning analytics. Especially for the CAs, it is emphasized that the latest studies systematically indicate that agent interventions during peer interaction (i.e. online discussions) trigger task relevant cognitive activity that leads to improved learning outcomes at many levels (domain-specific and domain-independent) [14][15][16]. The agent interventions are modeled according to what is known as "Academically Productive Talk" (APT) framework, which essentially refers to modeling the experienced teacher's "moves" (interventions) during students' dialogue to make students elaborate in the domain [17]. The learning experience that this type of agent provides to students is similar to having one more partner in their group trying to respond to this partner's prompts.

The strategic goal of the project is to make a significant contribution towards European universities in their development towards knowledge service providers to

the economy and society at large and to target the now transversal competencies of programming, computational and design thinking as well as concrete application of new technologies for teaching. While the EU universities jointly offer a high level of education, they trail behind in offering novel and transversal competencies as well as in developing joint and novel approaches to MOOC education.

Ultimately, the project aims to provide novel methods for teaching and learning to address the problem of attrition in online learning. These goals will be achieved by pursuing the following objectives:

- Develop new learning and teaching methods for MOOCs building on novel technologies in collaborative learning like conversational agents for MOOCs that are capable of analysing learner interaction, facilitate learners' self-regulation and –assessment, and that can be integrated in a MOOC platform.
- To promote innovative solutions to current and future challenges and for sustainable impact on Europe's education and training systems by granting open access to the conversational agents built within the project.
- To demonstrate and validate the built capacity for innovative teaching and learning methods and mainstreaming them to the existing education and training systems, by the design, execution and assessment of several pilots that orchestrate individual and collaborative learning activities.
- Spread the best practice pilots built in the project to all participating universities and other economic and societal stakeholders (educational authorities and civil society organizations for further training for the unemployed, businesses for further training of the workforce) as well as granting access to the MOOCs on European and national providers for free, openly accessible MOOCs, e.g., platform.europeanmoocs.eu.

The application of the results in multiple countries and large number of users will collect and analyze substantive evidence on our approaches. The usage of a well-established pedagogical model and the gathering, processing and analysis of the generated data with novel LA tools will provide informative evidence that can set the foundations for future concrete methodologies. These methodologies will foster participatory learning in a structured environment where learners will be able to develop transversal as well as domain-specific skills, becoming equipped for their future employability.

Finally, the colMOOC project will enforce market stakeholders in educational technology sector with new tools and methods.

3.2 Work plan

In order to achieve the above aims and objectives, the colMOOC project methodology will be divided into three major phases:
- Analysis & Design: During this phase all the possible user requirements are to be identified.
- Implementation: This phase regards the preparation and practical application of the designed results in real world settings and large number of end-users to test their validity and ability to enhance the quality of education and training.

- Evidence-based Policy Formation: This phase includes the accumulation, study and analysis of the evaluation results after the trials and the corresponding design of policy recommendations. These recommendations will be concrete and in accordance to the policy making lifecycle.

The project work has been divided in 6 operative work packages (WP) as follows:
- WP1: Educational approach construction and framework. This WP constitutes the needs analysis of the project and the user requirements in order to construct our educational approach.
- WP2: Conversational agent component. This WP constitutes the design and the development of the conversational agent component
- WP3: Learning analytics component. This WP constitutes the design and the development of the Learning analytics component
- WP4: System Integration. This WP refers to the integration of the previous components into several MOOCs platforms
- WP5: Courses preparation, trials and evaluation. This WP refers to the preparation for the testing of the results in real world educational and training conditions. Educators and trainers will design their courses according to the approach constructed during WP1. WP5 will develop exemplary interdisciplinary MOOC content for university students and specific workforces. Moreover, WP5 work will proceed with the actual trials that will test the results in large numbers of users and evaluate the process.
- WP6: Policy recommendations and Community of Practice. WP6 will start with the development of a Community of Practice platform, which will host all cooperation activities, project results and act as a community building space for colMOOC practitioners.

In addition to these 6 WPs, additional activities will include the dissemination, exploitation and impact of the project results (WP7) and the project management (WP8). A relevant visualization of the colMOOC workpackage main activities is presented in Fig. 1 below.
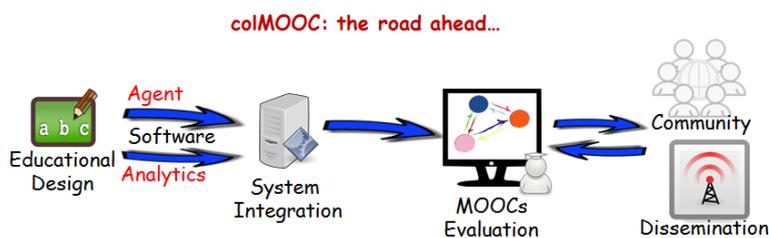
Fig. 1. colMOOC project main activities

Careful scheduling of WPs has been performed so that implementation of WPs run partly in parallel, so that each WP feeds initial input to the next WP, while the first WP has enough time to finalize its results according to feedback.

The relevant milestones (M) as the corresponding main success indicators along with the major project results are:

- M1. colMOOC educational approach and component design.
- M2. Components models and tools and first version of integration mechanism released.
- M3. Software modules configured, courses design and materials developed
- M4. Trials and evaluation completed. Validation activities in educational and training settings are performed and evaluated.
- M5. Policy recommendations drafted.

In terms of work methodology, the project will follow a four- step loop approach where education and training policies are used as input in order to produce the project's objectives and then provide new policy actions as outputs based on validated evidence. This way, the project's potential will be maximized in successfully transferring the project results into European policy development. These steps include:

- Study of good practice projects under various projects that have dealt with and successfully raised awareness and translated results into policy actions.
- In-depth study or existing models for policy development and the education and training policy agendas. This will include study of the nature of policy and the particular nature of education policy within its wider social, political and economic contexts.
- Design of innovative approaches and test of their validity on large-scale multilingual and multidisciplinary settings.
- In-depth study of wider policy recommendations in a national and/or European level for the identification of best practices.

Finally, the evaluation plan of the project aims to develop three pilot MOOCs (WP5 activities):

1. *Programming for Non-Programmers.* Programming skills are not only a major asset for getting hired for developing software, but also allow for better understanding and co-designing with a development team. Programming skills are required to use, build, and maintain digital infrastructures for managing and compiling data.
2. *Computational and Design Thinking.* This pedagogical strategy fosters systematic and creative approaches to problem solving. This involves analysis and formalization of complex problems, understanding and developing feasible, technical solutions, as well as divergent and convergent ways of thinking.
3. *Educational Technologies in the Classroom.* New technologies enhance learning through dynamic, interactive, multi-media formats in simulations, web-videos, and games. Advanced technologies enable new forms to aggregate and visualize data for feedback to students and teachers.

The above project activities lead to the preparation of input state-of the art papers (dissemination activities within WP7) that provide recommendations for the formulation of new policy actions in the university-level and training-level.

4. Results

In this section, the expected results of the project will be first shown and then how these results will benefit each of the identified target groups will be explained.

4.1 Expected results

The results of the colMOOC project are expected to benefit the higher education institutions in different ways:

- Enhanced MOOC courses: the project will provide prototype MOOC courses and services that will increase students' online engagement in conditions of productive dialogue and thus deeper learning. This, in turn, is expected to increase students' motivation and interest to follow the course. Moreover, university teachers will be able to use new and innovative pedagogical approaches such as collaborative learning supported by conversational agents to produce similar MOOCs for a diverse, larger audience.
- Minimize the dropout rate by integrating learning analytics based students' assessment and the cost of accreditation.
- As a marketing tool, increase the student enrolment by offering short MOOCs of different disciplines to potential students of formal academic programs, where they can see, try and understand the university pedagogical model and technological advances before the formal enrolment.
- New research opportunities.

  Moreover, enterprises are also going to be benefitted from the project outcomes in two main ways:
- Enforce their market strategy with new tools and methods
- New market areas for not educational technology ICT companies

These results cannot be achieved without cooperation at national, regional or local level for the following reasons:
- The pilots will be implemented in four different countries (Greece, Spain, Germany, Denmark)
- In order to provide a more holistic approach that will be available to the market, the proposed solution should be able to be applied in different educational domains and approaches
- The consortium includes experts in all fields (experts in pedagogy, MOOCs, conversational agents, learning analytics and educational scenarios).

4.2 Target groups

The project targets three groups: the academic community, the training community, and the technology providers. As these groups transcend the project consortium, the project targets people through three networks. The first network is comprised by the project partners (e.g., researchers, educators, administrative staff, management of

partner organizations), the second by the immediate networks of each of the project partners (e.g., municipalities, school/university networks, enterprises, associations), and the third by all other actors interested in MOOCs that could reach the project through the online presence and the outreaching activities planned in the project. Next, the benefits from the project results for each of these groups are explained:

Academic community
- University students. Enrolled students require increased flexibility and support, something that cannot be always available in a university system that deals with rising number of students and limited funds.
- University teachers. As MOOC elements and online pedagogies are integrated into university curricula, the use of techniques, such as learning analytics, could provide a holistic view of the learning activity would be valuable for the teachers, in their effort to improve their courses.
- Academic institutions. The integration of conversational agents and learning analytics into MOOCs would allow the universities to offer more flexible and engaging ways of teaching. This would allow the universities to reach wider demographic (based on age, occupation, etc.) and address the need for lifelong learning.
- Researchers. Cutting-edge original research will be explored during the project. To the best of our knowledge there are only two research groups currently worldwide developing this kind of agents with APT based dialogue modeling and interventions, one of these groups participating in the project [14].

Training community
- Learners. The project will provide an enhanced learning experience for the online learner, offering at the same time different ways to formal education, lifelong learning, and training.
- Trainers. Trainers are in need of innovative, engaging and motivational learning methods in order to attract the attention of the employees of any sector during training sessions. In addition, trainers should be able to easily design courses and material and monitor their trainees' progress during sessions. This way, they will be able to adapt the process accordingly, change the real world scenarios used as problems that need solving and foster the employees' skills development.
- SMEs. European SMEs continuously need to train their employees to improve the competitive advantage of the enterprise. SMEs can be benefited by adopting a colMOOC training approach thus being able to provide personalized training to their employees and being able to monitor their progress.

Technology providers
- Tool developers. Learning tool developers (e.g., MOOC plug-in developers) need to provide high quality personalized software products to their customers.
- Content providers. Learning content providers (e.g. MOOC providers) aim at making their content easily accessible to a great number of users.

Eventually, the project target results are in accordance to the main goals of the European policy agendas as this will enable and facilitate the dialogue between the consortium and the policy makers as well as their transfer to actual policy development.

5 Conclusions and ongoing work

<mark>To be completed.</mark>

References

1. Lankshear, C., & Knobel, M. (2008) (Eds.). *Digital literacies: concepts, policies and practices*. New York: Peter Lang. ISBN 9781433101687.
2. Siemens, G. (2013). Massive open online courses: Innovation in education. *Open educational resources: Innovation, research and practice*, 5.
3. Tegos, S., Demetriadis, S., & Karakostas, A. (2015). Promoting academically productive talk with conversational agent interventions in collaborative learning settings. *Computers & Education*, 87, 309-325.
4. Bassi, R., Daradoumis, T., Xhafa, F., Caballé, S., Sula, A. (2014). Software Agents in Large Scale Open E-learning: A Critical Component for the Future of Massive Online Courses (MOOCs). *In proceedings of the Sixth IEEE International Conference on Intelligent Networking and Collaborative Systems*, pp. 184-188. IEEE Computer Society.
5. Daradoumis, T., Bassi, R., Xhafa, F., Caballé, S. (2013). A review on massive e-learning (MOOC) design, delivery and assessment. *In proceedings of the Eighth International Conference on P2P, Parallel, Grid, Cloud and Internet Computing*, pp. 208-213. IEEE Computer Society.
6. Barak, M., Watted, A., & Haick, H. (2016). Motivation to learn in massive open online courses: Examining aspects of language and social engagement. *Computers & Education* 94, 49-60
7. Capuano, N., Caballé, S. (2015) Towards Adaptive Peer Assessment for MOOCs. *In proceedings of the 10th International Conference on P2P, Parallel, Grid, Cloud and Internet Computing*, pp. 64-69. IEEE Computer Society.
8. Schuwer, R., Jaurena, I. G., Aydin, C. H., Costello, E., Dalsgaard, C., Brown, M., & Teixeira, A. (2015). Opportunities and threats of the MOOC movement for higher education: the European perspective. *The International Review of Research in Open and Distributed Learning*, 16(6).
9. Daradoumis, T., Bassi, R., Xhafa, F., Caballé, S. (2013). A review on massive e-learning (MOOC) design, delivery and assessment. *In proceedings of the Eighth International*

Conference on P2P, Parallel, Grid, Cloud and Internet Computing, 208-213. IEEE Computer Society.

10. Miguel, J., Caballé, S., Prieto, J. (2013). Providing Information Security to MOOC: Towards effective student authentication. *In proceedings of the Fifth IEEE International Conference on Intelligent Networking and Collaborative Systems*, 289-292. IEEE Computer Society

11. Capuano, N., Caballé, S., Miguel, J. (2016). Improving Peer Grading Reliability with Graph Mining Techniques. Editorial. *International Journal of Emerging Technologies in Learning*, 11(7), 24-33.

12. Miguel, J., Caballé, S., Prieto, J. (2013). Providing Information Security to MOOC: Towards effective student authentication. *In proceedings of the Fifth IEEE International Conference on Intelligent Networking and Collaborative Systems*, 289-292. IEEE Computer Society.

13. Gañán, D., Caballé, S., Clarisó, Conesa, J., Bañeres, D. (2017). ICT-FLAG: A Web-based e-Assessment Platform Featuring Learning Analytics and Gamification. *Journal of Web Information Systems* 13(1), 25-54

14. Tegos, S., Demetriadis, S., & Karakostas, A. (2015). Promoting academically productive talk with conversational agent interventions in collaborative learning settings. *Computers & Education*, 87, 309-325.

15. Karakostas, A., & Demetriadis, S. (2014). Adaptive vs. Fixed Domain Support in the Context of Scripted Collaborative Learning. *Educational Technology & Society*, 17 (1), 206–217.

16. Tegos, S., Demetriadis, S., & Tsiatsos, Th. (2013). A Configurable Conversational Agent to Trigger Students' Productive Dialogue: A Pilot Study in the CALL Domain. *International Journal of Artificial Intelligence in Education*. *24*(1), pp. 62-91. Springer: New York. DOI: 10.1007/s40593-013-0007-3

17. Kumar, R., & Rose, C. P. (2011). Architecture for Building Conversational Agents that Support Collaborative Learning. *IEEE Transactions on Learning Technologies*, 4(1), 21-34.

18. Michaels, S., O'Connor,M. C., Hall, M.W., & Resnick, L. B. (2010). Accountable talk sourcebook: For classroom that works (v.3.1). University of Pittsburgh Institute for Learning. Retrieved March 1, 2017, from https://www.ortingschools.org/cms/lib/ WA01919463/Centricity/domain/326/purpose/research/accountable%20sourcebook.pdf

19. Noroozi, O., Weinberger, A., Biemans, H. J. A., Mulder, M., & Chizari, M. (2013). Facilitating argumentative knowledge construction through a transactive discussion script in CSCL. *Computers and Education*, 61(2), 59–76.