Preface

Online education and especially massive open online courses (MOOCs) 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. MOOCs have been reported as an efficient and important educational tool, yet there are a number of issues and problems related to their educational impact. More specifically, there is an important number of dropouts 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.

Previous studies combine Artificial Intelligence (AI)-based approaches, such as the use of conversational agents (CA), chatbots, and data analytics in order to face the above challenges. However, these studies explore these and other AI approaches separately, thus having less impact in the learning process. Therefore the effective integration of AI novel approaches in education in terms of pedagogical CA and learning analytics (LA) will create beneficial synergies to relevant learning dimensions, resulting in students' greater participation and performance while lowering dropout rates and improving satisfaction and retention levels. In addition, tutors, academic coordinators, and managers will be provided with tools that will facilitate the formative and monitoring processes.

Specifically, the book aims to provide novel AI and analytics-based methods to improve online teaching and learning, addressing key problems such as the problem of attrition in MOOCs and online learning in general. To this end, the book contributes to the educational sector at different levels:

- Deliver new learning and teaching methods for online learning (with a specific focus on MOOCs), building on novel technologies in collaborative learning, such as CA and LA, that are capable of boosting learner interaction and facilitate learners' self-regulation and -assessment.
- Demonstrate and validate the built capacity for innovative teaching and learning methods and mainstream them to the existing education and training systems, by the

design, execution, and assessment of pilots that orchestrate individual and collaborative learning activities.

- Promote highly innovative solutions and beyond the state-of-the-art models for online and MOOC-based learning and implementations with the integration of AI services, such as, for example, based on CA and LA, to face current and future challenges and for sustainable impact on online educational and training systems.
- Demonstrate and exemplify efficient teaching techniques leveraging the power of analyzing data generated by smart AI-based interfaces, such as those promoting interactions with CA in learning environments.
- Deepen our understanding of how CA tools can contribute to increasing the transactional quality of peers' dialogue and, consequently, the quality of learning, in various situations, such as learning in academic settings and also corporate training in business environments.

The ultimate aim of this book is to stimulate research from both theoretical and practical views, including experiences with open source tools, which will allow other educational institutions and organizations to apply, evaluate, and reproduce the book's contributions. Industry and academic researchers, professionals, and practitioners can leverage the experiences and ideas found in the book.

This book consists of 15 chapters, starting with the Introductory Chapter where the Book Editors, led by Stavros Demetriadis, present the European project "colMOOC," which supports the edition of this book. The aim of this leading chapter is to describe the rationale of the project motivated by the issues found in the context of MOOCs, which provide a powerful means for informal online learning that is already popular, engaging great numbers of students all over the world. However, studies on MOOCs efficiency frequently report on the high dropout rates of enrolled students, and the lack of productive social interaction to promote the quality of MOOC-based learning. The project proposes and develops an agent-based tool and methodology for integrating flexible and teacher-configurable CA along with relevant LA services in online educational platforms, aiming to promote peer learning interactions. The authors claim that CAs appear to be a promising AI technology with the potential of acting as catalysts of students' social interaction, a factor known to beneficially affect learning at many levels. From this perspective, the chapter provides reflections on the first project outcomes emerging from four different pilot MOOCs. Early conclusions analyze the challenges for integrating a teacher-configured CA-chat service in MOOCs, provide helpful guidelines for efficient task design, and highlight promising evidence on the learning impact of participating in agent-chat activities.

The rest of the book chapters are organized into three major areas:

Part I: *Intelligent Agent Systems and Learning Data Analytics*: The chapters in this area address the use of pedagogical CAs and LA to provide supportive, personalized,

and interactive online teaching and learning in learning management systems (LMSs) and in particular in massive education as in MOOC platforms. Benefits and challenges of the proposed educational strategies supported by these technological approaches are unveiled and the research results are illustrated with practical adoptions in real contexts of learning. The cross-cutting scope of the research approaches can be applied to different knowledge areas and learning modes and styles with the ultimate purpose to improve and enhance the online teaching and learning experience.

Part II: *Artificial Intelligence Systems in Online Education*: This area starts with exploring the intersection between AI, online education, and ethics with the aim to draw people's attention to the ethical concerns surrounding this crossroads. The rest of the chapters in the area provide AI-based solutions to address relevant issues found in current online education, such as poor personalization, high academic dropout, learners' disengagement, and low participation, many of them resulting from facing online education at scale and big data. To deal with these issues, the chapters propose to use different AI techniques, such as machine learning, sentiment analysis, and natural language processing. Simulation results in terms of technical performance and accuracy are compared with similar approaches, and implications of these results for online education are illustrated in terms of improving the effectiveness of the online teaching and learning process at scale.

Part III: *Applications of Intelligent Systems for Online Education*: The chapters covering this area present the applicability of different approaches of intelligent learning systems to various domains and for a variety of purposes, namely the analysis of socioemotional profiles within educational groups, to overcome the uniformity of online learning contents to deal with heterogeneous learners, to support learners with varying reading difficulties, and to improve teaching and learning of science, technology, engineering, and mathematics (STEM) subjects. Strong implications and further challenges of the application of these approaches include making online education more effective, multidisciplinary, and collaborative, personalized, and fair.

The chapters in the first area of *Intelligent Agent Systems and Learning Data Analytics* are organized as follows:

Atif et al. in Chapter 1, AI-Enabled Remote Learning and Teaching Using Pedagogical Conversational Agents and Learning Analytics, claim that the advancements in AI have potentially created new ways to teach and learn, such as the use of LA to monitor and support students using data captured in LMSs. To back up this claim, the authors in the chapter report the benefits of using AI-enabled CAs in multiple units/subjects across two universities and illustrate how these CAs can play a role similar to a teacher or peer learner by sharing the expertise they have acquired from the knowledge contained in student−teacher social interactions in LMS forums and grade-book teacher feedback.

The chapter shows how unlike teachers or peers, these CAs can be contacted anonymously at any time, they do not mind being asked the same question repeatedly, and they can empower students to explore options and outcomes. The chapter concludes with a discussion of the potential of LA to automate CA interactions.

Chapter 2, Integrating a Conversational Pedagogical Agent into the Instructional Activities of a Massive Open Online Course, by Rizzardini et al. addresses the topic of using pedagogical CAs to offer a wide range of possibilities when incorporated into virtual training courses. The chapter is motivated in the context of MOOCs, where the interaction with the students is at scale, thus hindering personalized interaction by human teachers. The authors believe that an adequate configuration of pedagogical CAs has the potential to provide personalized attention. However, the authors claim that there are no "one-size-fits-all" approaches in terms of pedagogical CAs given that the conversations usually start from scratch, without much user context, becoming especially problematic when addressing the issue of scalability in MOOCs where students show different states and a similar approach is not useful for all of them, requiring to start with a previous context. To address this issue, the authors propose the use of LA to provide a better context for decision-making and initial values to launch the model resulting in greater possibility of success. To this end, the goal of the chapter is to present a prototype integrating a CA embedded into the instructional activities of a MOOC with the ultimate aim to increase motivation and student engagement to achieve their learning goals while producing improvements in students' behavior and higher completion rates.

Abdullah and Sakr in Chapter 3, Improving MOOCs Experience Using Learning Analytics and Intelligent Conversational Agent, discuss the effectiveness that online learning has proved in the last years among a wide range of learners. In particular, the authors claim that MOOCs have revolutionized the shape of learning as a substitutional tool compared to the conventional educational system, due to, among other reasons, their flexibility in timing, elimination of economic and geographical constraints, while enabling learners from different cultures to communicate and share their knowledge through online forums. Then, the authors turn the discussion into the challenges found in MOOCs that need to be faced, such as higher dropouts rates among learners at different phases of the course and reduction in participation level of learners. The chapter aims to address these challenges while enhancing the MOOCs experience through the provision of an innovative framework named Learning Analytics Technique and Intelligent Conversational Agent (LAICA) with the purpose of integrating LA and intelligent CAs to improve the MOOC experience for learners and educators. The chapter provides a throughout description of the LAICA framework from the architectural view, and a case study of implementation and integration of the framework in a MOOC is provided.

In Chapter 4, Sequential Engagement-based Online Learning Analytics and Prediction, Song and Li analyze how online education has become a widely accepted teaching method

over the recent years as an integrated learning platform providing learning materials and assessment tools. In their analysis, the authors claim that through the complete access rights and records of the students' complete activities on the learning platform, students' learning engagement and evaluation results can be well analyzed and predicted. However, with the development and changes in teaching contents, the authors claim that new challenges have emerged as various new forms of textbooks and interactive methods have been introduced into various online education platforms, which make more implicit learning patterns be learned, resulting in students' online activities being closely related to their final grades. From this motivation, the authors simulate learning activities in a new teaching format in order to accurately predict their final performance by leveraging important research outcomes in the LA field. Eventually, the chapter aims to explain in detail how to integrate the latest LA research methods into modeling students' sequential learning engagement, so as to accurately predict students' learning performance.

Kuk et al. concludes the first area of the book in Chapter 5, An Intelligent System to Support Planning Interactive Learning Segments in Online Education, by discussing the intelligent educational services used by today's LMS platforms for the purpose of creating personalized learning environments. The authors claim that these learning environments must be adapted for personal student leaning style. To this end, the purpose of the chapter is to propose, develop, and explain the implementation of a personalized intelligent system. The proposed system suggests additional learning resources that will support students' immersive learning process, which will lead toward better outcomes of learning activities. However, the authors consider that online e-learning systems should implement successful methods and evaluation techniques when taking different teaching paths though facing technical challenges still unsearched, which are the main motivation of the chapter. To this end, the authors in the chapter urge that every e-learning system should have different interactive learning segments in the form of learning objects in text, video, image, quiz, etc., as entities in each separate course in the e-learning system, and supported by LA techniques as the most appropriate method for automatic detection of student learning models. To this end, the chapter presents LA techniques for analyzing learning paths composed from four different learning objects, which are then implemented in a Moodle environment. As a result, generalized sequence patterns are mapped, and an activity module named Observer is used to track students' learning behavior. The results of the tracking are eventually used to develop an intelligent system for planning interactive learning segments.

The chapters in the second area of *Artificial Intelligence Systems in Online Education* are organized as follows:

Chapter 6, A Literature Review on Artificial Intelligence and Ethics in Online Learning, by Casas-Roma and Conesa draws attention to how AI is being used in online learning to improve teaching and learning, with the aim of providing a more efficient, purposeful,

adaptive, ubiquitous, and fair learning experiences. However, the authors claim that, as it has been seen in other contexts, the integration of AI in online learning can have unforeseen consequences with detrimental effects that can result in unfair and discriminatory educational decisions. Therefore the main authors' motivation is that it is worth thinking about potential risks that learning environments integrating AI systems might pose. To this end, the authors explore the intersections between AI, online education, and ethics in order to understand the ethical concerns surrounding this crossroads. As a result, the chapter provides an extensive review work on the main ethical challenges in online education identified in the literature while distilling a set of guidelines to support the ethical design and integration of AI systems in online learning environments. The authors conclude that the proposed guidelines should help to ensure that online education is how is meant to be: accessible, inclusive, fair, and beneficial to society.

Capuano in Chapter 7, Transfer Learning Techniques for Cross-domain MOOC Forum Postanalysis, addresses the role of discussion forums, as popular tools in the context of MOOCs, used by students to express feelings, exchange ideas, and ask for help. Due to the high number of students enrolled and the small number of teachers, the author claims that the automatic analysis of forum posts can help instructors to capture the most relevant information for moderating and carefully planning their interventions. To this end, the author first explores several emerging approaches to the automatic categorization of MOOC forum posts and claims that such approaches have a common drawback given that when they are trained on labeled forum posts from one course or domain, their application on another course or domain is often unsatisfactory. For instance, different courses have different feature spaces and distributions, and certain words may appear frequently in one course, but only rarely in others. To help overcome this drawback, the author then introduces a cross-domain corpus-based text categorization tool that includes transfer learning capabilities for the detection of intent, sentiment polarity, level of confusion, and urgency of MOOC forum posts. The underlying model, based on convolutional and recurrent neural networks, is trained on a standard labeled dataset and then adapted to a target course by tuning the model on a small set of labeled samples. The proposed tool reported in the chapter is eventually experimented with and compared with related works.

Chapter 8, Assisted Education: Using Predictive Model to Avoid School Dropout in E-Learning Systems, by Neves et al. discusses the important issue of students dropping out of school as a real challenge for educational specialists, especially in distance education classes, which deal with a huge number of students' disengagement with social and economic costs. In this context, the authors claim that behavioral, cognitive, and demographic factors may be associated with early school dropout. Motivated by this claim, the aim of the chapter is to propose an enhanced machine learning ensemble predictive architecture capable of predicting the disengagement of students along with the class. The system notifies teachers, enabling them to intervene effectively and make students' success

possible, and students to give them a chance to turn back. To evaluate the proposed architecture, the chapter provides a case study showing the feasibility of the solution and the use of its technologies. Evaluation results point out a significant increase of gain in accuracy along with the class, reaching a high level of precision.

Azoulay et al. in Chapter 9, Adaptive Task Selection in Automated Educational Software: A Comparative Study, consider the challenge of adapting the difficulty level of the tasks suggested to a student using an educational software system. In their study, the authors investigate the effectiveness of different learning algorithms for the challenge of adapting the difficulty of the tasks to a student's level and compared their efficiency by means of simulation with virtual students. According to the results, the authors demonstrate that the methods based on Bayesian inference outperformed most of the other methods, while in dynamic improvement domains the item response theory method reached the best results. Given the fact that correctly adapting the tasks to the individual learners' abilities can help them increase their improvement and satisfaction, this chapter can assist the designers of intelligent tutoring systems in selecting an appropriate adaptation method, given the needs and goals of the educational system, and given the characteristics of the learners.

Chapter 10, Actor's Knowledge Massive Identification in the Learning Management System, by Touimi et al. concludes the second area of the book by discussing on the generation of traces in any computer system either by user interactions with the system or by the system itself. The authors claim that with the proliferation of new technologies, computer traces keep increasing, rapidly and brutally making enormous changes in the field of education in terms of technical means and teaching pedagogy. In the context of online education, the emergence of MOOCs offers unlimited free access over time and space where interactions by learners generate large amounts of data that are difficult for tutors and learners to process in learning platforms. The authors focus on the need for learners to build, share, and seek knowledge in a MOOC through discussion forums, which are an efficient tool for communication, sharing ideas, opinions, and seeking answers to learners' questions. As a contribution to this research field, the aim of the chapter is to report the development of a framework capable of managing big data in discussion forums in order to extract and present relevant knowledge, which is crucial in the case of MOOCs. The framework is based on the process of analyzing learners' trace log files, which includes the stages of collection, statistical analysis, and then semantic analysis of traces of learners' interactions. As statistical analysis reduces the dimensionality of the data and builds new variables, the authors propose the Latent Dirichlet Allocation Bayesian inference method be applied to threads and messages posted in the discussion forums in order to classify the relevant response messages, present a semantic response to the learners, and enrich the domain ontology with new concepts and new relationships. The framework uses the Apache Spark libraries for the computation speed constraints.

The chapters in the third and last area of *Applications of Intelligent Systems for Online Education* are organized as follows:

Fotopoulou et al. in Chapter 11, Assessing Students' Social and Emotional Competencies Through Graph Analysis of Emotional-Enriched Sociograms, address the topic of social and emotional competencies of students and discuss whether the improvement of these competences is associated with positive effects in various learning, collaboration, and personal development activities. The discussion is focused on social and emotional learning activities, which are developed and applied by tutors within educational groups and the effective application of such activities requires knowledge regarding the status and evolution of socioemotional profiles at the individual and group level. In this context, the authors claim that graph theories can be proven helpful for supporting the monitoring and analysis of the evolving relationships of socioemotional characteristics, considering the interactions among group members. As a response to these claims, the chapter proposes a methodological approach for the composition and analysis of social and emotional profiles of individuals within educational groups, based on the creation of socioemotional graphs. Upon the realization of a literature review over sociometric assessment and emotional competencies representation and development techniques, the chapter details the proposed methodological approach that focuses on the joint creation and analysis of socioemotional profiles within educational groups. Applicability of the approach in various domains is presented, along with the description of identified challenges that can be tackled in the future.

Chapter 12, Intelligent Distance Learning Framework for Heterogeneous Online Students: Assessment-Driven Approach, by Yoon discusses the role of online education for various remote training, including their inevitable adoption in anomalous situations (e.g., the COVID-19 pandemic). The discussion is contextualized in MOOCs as well-recognized online education platforms whose contents are developed by world-famous experts. However, the author claims that the learning effectiveness of online education is not yet higher than face-to-face classroom educations, in part because the online content is uniformly designed for heterogeneous online students. Motivated by this claim, the author proposes both online lecture and mobile assessment platforms to elevate the quality of distance learning. The online platform proposed has three layers (namely, basic, advanced, and application) per course module with each layer divided into four quadrant panes (namely, slides, videos, summary, and quizlet). Students begin with the basic layer first, dive into the advanced layer, and then move on to the application layer. In addition, mobile assessment platforms are proposed to face the critical issue in distance learning of fair assessment management. In particular, the author proposes smartphone-based assessment of online student learning performance to disable the high chance of cheating schemes, and to enable the building of student learning patterns as the analytics will eventually lead to the reorganization of online course module sequences. Upon these considerations, the main

contributions of the chapter are: (1) an accurate recognition of student weakness; (2) an intelligent and automatic answering to student questions; and (3) a mobile phone application-based assessment.

Jeske et al. in Chapter 13, Personalizing Alternatives for Diverse Learner Groups: Readability Tools, address the challenge for educators to identify new ways to meet the needs of increasingly diverse educational cohorts. This conceptual chapter proposes that a number of technological developments (LA, intelligent tutoring systems) may set the stage for the incorporation of new tools—tools that can support learners with varying reading difficulties, from those who are not native speakers of the language of instruction to those who struggle due to a learning disability. The authors make suggestions regarding the general steps which would be required to implement existing readability tools into tutoring and learning management systems effectively. Particular attention is paid in the chapter to the concept of personalization of learning and the need for multidisciplinary collaboration to meet diverse learners' support needs.

The third and last area of the book is concluded in Chapter 14, Human Computation for Learning and Teaching or Collaborative Tracking of Learners' Misconceptions, by Heller and Bry who propose a technology-enhanced learning application which exploits both human computation and LA to improve the learning and teaching of STEM while reducing the teachers' workload. Specifically, the application aims at tracking and reducing learners' misconceptions and easing teachers' corrections of homework. To achieve this goal, the proposed software uses text processing, collaborative filtering, and teacher collaboration in a wiki-like environment. While the software has not been implemented as a whole, its main components have been implemented and evaluated in several case studies which are reported in the chapter. These studies point to the realizability and effectiveness of the approach.

Final words

Through 15 selected quality chapters the book covers scientific and technical research perspectives that contribute to the advance of the state-of-the-art whilst providing a better understanding of the different problems and challenges of current online education. In particular, the book proposes a great variety of innovative solutions by AI-based methods and techniques with a broad focus on the use of pedagogical CAs and LA in MOOCs environments to ultimately improve the effectiveness of the online teaching and learning. The applicability of the proposed strategies and solutions are exposed in the book along with strong implications in terms of benefits and challenges illustrated by detailed case studies in real contexts of online education.

Researchers will find in this book the latest trends in these research topics. Academics will find practical insights into how to use conceptual, practical, and experimental approaches in their daily tasks. Meanwhile, developers from the online education community can be inspired and put into practice the proposed models, methodologies, and developments to evaluate them for specific purposes within their own application context.

Finally, we would like to thank the authors of the chapters and also the referees for their invaluable collaboration and prompt responses to our enquiries, which enabled the completion of this book in due time. We also thank Prof. Yannis Dimitriadis for his excellent contribution to the foreword of this book. Last, but not least, we gratefully acknowledge the feedback, assistance, and encouragement received from the Editor-in-Chief of this Elsevier Book Series, Prof. Fatos Xhafa, and Elsevier's Project Manager, Andrea Dulberger.

We hope the readers of this book will find it a valuable resource for their research, development, and educational activities.

<div align="right">

Santi Caballé, Stavros N. Demetriadis, Eduardo Gómez-Sánchez,
Pantelis M. Papadopoulos and Armin Weinberger

</div>