When E-Learning Meets Big Data, Cognitive Computing, and Collaborative Environments

Mauro Coccoli

University of Genoa
DIBRIS
Genoa, Italy
mauro.coccoli@unige.it

Paolo Maresca

Federico II University
DIETI
Naples, Italy
paolo.maresca@unina.it

Andrea Molinari

Dept. of Industrial Engineering
University of Trento, Italy
School of Industrial Engineering and
Management, Lappeenranta
University of Technology, Finland
andrea.molinari@unitn.it

Abstract— The e-learning scenario is fast-changing in all of its main aspects from both the technology side and the users' perspective. In this vein, applications and services as well as methodologies are evolving rapidly running after the more recent innovations, through the adoption of novel infrastructures and distributed architectures, to provide the most advanced solutions. In this situation, we observe that the more influential technological factors are two. The former is the availability of a new wave of big data also inherent educational tasks, which can be considered a powerful source for all the people involved in the educational process and acts as enabler for the development of new models for both management and education, based on analytics. The latter is cognitive computing, which can provide learners and teachers with innovative services enhancing the whole learning process, also introducing new capabilities to improve human machine interactions. In these circumstances, it seems that Learning Management Systems (LMS) have reached their maturity stage and they may need a redesign. In this respect, the paper outlines the evolutionary trends of Technology Enhanced Learning (TEL) and presents the results achieved within two experiences carried on in two Italian universities.

Keywords-cognitive computing; big data; collaborative systems; technology enhanced learning; e-learning

I. INTRODUCTION

Both technology and society are undergoing a continuous evolution process, which is pushing innovation and driving the development of novel solutions in almost any field of application. The wide adoption of such solutions in a plethora of services that are commonly available on the Web for daily operations heavily impacts on users' behaviors and expectations. Specifically, with reference to the Technology Enhanced Learning (TEL) scenario, we observe significant steps ahead in techniques and methodology; consequently, the relevant technological solutions are subjected to continuous upgrades to cope with these, to the aim of improving the quality of services, the usability, the overall performances, the effectiveness of education, and to provide learners with a more pervasive experience. Accordingly, e-learning environments and the related tools have been growing in complexity, i.e., traditional Learning Management Systems (LMS), based on a centralized software architecture, are moving towards clusters of Massive

Open Online Courses (MOOC) platforms in cloud-based distributed architectures [1] and Learning Objects (LO) include videos and multimedia interactive artifacts. Concurrently, LMSs have been exploring new spaces of possibilities; among these, we mention the adoption of users' profiling techniques and analytics to the aim of tailoring personalized learning paths and activities as well as to predict students' careers, to achieve a further empowerment of the individual learning model. Anshary et al. provide us with a clear synthesis of the more significant evolutionary steps of learning solutions [2] by depicting a timeline that shows the changes occurred in tools, services, learning strategies, methodologies, and delivery techniques. The authors highlight that each progress achieved is not a replacement of the previously adopted solution, yet it implements one or more of the following: *(i)* extend capabilities; *(ii)* improve performances; *(iii)* promote different educational approaches and methodologies; *(iv)* change the way contents are delivered; *(v)* change users interactions model. Despite the methodology is sound, their analysis concerns the period from 2001 to 2015 and, since then, many things have evolved mainly owing to the rise of big data and cloud computing.

In this respect, we highlight that the adoption of cognitive computing solutions can be the technological enabler for a number of new functionalities [3]. Moreover, the introduction of the cognitive computing paradigm also impacts on the learning process, as schools and universities must face with new jobs and new training demands, where big data contexts are fundamental, since new knowledge and skills should be promptly delivered [4]. As a result, it seems that LMSs have reached their maturity stage in the innovation adoption curve. In fact, they are pervasively adopted by learning providers at any level, from the primary school to higher education institutions, to guarantee the effective implementation of TEL solutions, yet they hardly can cope with the high degree of interoperability and complexity that novel paradigms require, hence should expand their boundaries offering new services and also involving industries and the labor market in the learning design and process. In this respect, the paper discusses the evolutionary trends of e-learning and presents the results achieved within two experiences carried on in two Italian universities: *(i)* an innovative didactic approach to software engineering, which adopts advanced technologies and mixed solutions in a cloud infrastructure; *(ii)* a novel

collaborative learning environment that surpasses the typical functionalities of the more prominent LMSs. We highlight that such experiences are driving TEL platforms towards new standards where both learners and teachers, rather than technology, will be at the center of the process.

The remainder of the paper is structured as follows: in Section 2 we present some related works investigating these topics; Section 3 outlines the main drivers for the development of the next generation TEL platforms; in Section 4 we report the experience of a cognitive-computing-based laboratory, while in Section 5 we depict a newly developed e-learning collaborative environment. Section 6 concludes the paper also giving hints on future development directions.

II. RELATED WORKS

As highlighted in the previous section, we are going to focus mainly on two phenomena: *(a)* big data, and *(b)* cognitive computing. In fact, the possibility of gathering information from a wide number of heterogeneous sources, combined with the unprecedented opportunities of processing such data by means of cognitive computing techniques, is reshaping the technological scenario of e-learning applications, allowing the improvement of existing methodologies as well as imagining new ones. Moreover, the possibility of recording, storing and aggregating information, can significantly improve learning performances of both students and teachers [5]. For these topics, we report a number of related works that illustrate different solutions showing how variedly these two aspects can empower e-learning methodologies and how platforms are evolving with new functionalities. From a quick analysis of the relevant literature, i.e., searching indexing services such as Science Direct, Scopus and Web of Science, we found several publications on these challenging topics, enlightening that, more and more, LMSs offer such services among basic tools or as add-on functionalities.

A. The Impact of Big Data on E-Learning

Due to the presence of big data, various changes have occurred in educations and in e-learning systems. Their impact on the academic environment, which covers a large scope and has specific peculiarities, is investigated by Banica and Radulescu. Their expectation is that a change will suddenly occur in the way e-learning is approached by students and teachers. In this respect, based on the currently available software solutions, they propose a system architecture fostering universities to constitute consortia to analyze, organize and access huge data sets in common, in a cloud-based environment [6]. A similar approach is followed by Mahmod and Ali, which propose a strategy to improve the education process output through the collaboration with other international universities, to the aim of supporting the teaching and learning process by means of shared e-learning services. The integrated framework they propose links the individual impact with the organizational impact, promoting a collaborative culture model [7]. As hinted, another key factor for the empowerment of e-learning services, is their integration with information from other platforms and services given that a large amount of data is accessible in, e.g., online communities, blogs, discussion forum, messaging services, and social network sites. In this respect, Dietz-Uhler

and Hurn showcase some effective learning analytics techniques to derive knowledge from large blobs of information. Specifically, they focus on tracking students' data, to help them succeed. To this aim, they investigate which different learning analytics tools are adopted in different universities and institutions and discuss how faculty can exploit such data to monitor and predict students' performances, to finally enhance them [8]. Moreover, Kolekar et al. are observing the same large amount of information freely available over the Web, exploring the opportunities of using such data to get enhancements not only in assessments but in all the stages of the learning process [9]. Following the same current, Yu and Jo also show an example of how to fruitfully exploit big data for the prediction of students' performances, to the aim of optimizing their careers [10]. We point out that, currently, social media can play a vital role with respect to e-learning systems and the effectiveness of information is strongly tied to the way we process these data. In this respect, also Sheshasaayee and Malathi believe that the application of big data with e-learning is a hot topic, which has the potential for creating a huge impact on the whole education system [11].

We also highlight that processing big data in an effective way is possible only thanks to cognitive computing solutions implementing suited machine learning techniques, able to cope with data characterized by large volume, different types, high speed, uncertainty and incompleteness, and low value density. Going deep in technical details, algorithms and methods (e.g., representation learning, deep learning, distributed and parallel learning, transfer learning, active learning, and kernel-based learning) is out of the scope of this paper, however the interested readers can find a survey of how machine learning is used for big data processing in a paper by Qiu et al. [12]. Other references can be found in [13], where Gupta et al. present a large understanding of past, present and future directions in this domain, made through a mapping of the characteristics of cognitive computing, i.e., observation, interpretation, evaluation and decision, versus the so-called V's of big data, i.e., Volume, Variety, Veracity, Velocity and Value.

B. The Impact of Cognitive Computing on E-Learning

From the side of cognitive computing, only recent works can be found that deal with more advanced services for e-learning applications. For example, they can serve the instructional design process, helping to find personalized learning assets and improving the definition of individual learning strategies or classifying resources. Specifically, Leitão et al. focus on recommendation systems implemented on the basis of cognitive services. They envisage the possibility of using the same approach to meet the needs of students and teachers, especially to enable personalized learning strategies and recommend educational resources, based on information derived from the interactions between students. The paper proposes the prototype of a platform and surveys the approaches to develop advanced TEL solutions, analyzing the state of the art of using cognitive systems in e-learning, identifying paradigms and pedagogical methodologies, techniques, tools and learning objects with respect to the recommendation of pedagogical activities using cognitive computing [14]. In [15], the authors use cognitive systems for the automated classification of learning videos, with special reference to MOOCs. Dessi et al. exploit the capability

of Automatic Speech Recognition (ASR) and Optical Character Recognition (OCR) services to extract text from audio and visual frames so to be able to perform a classification based on a taxonomy. Further development of their work led the authors to find a solution to analyze the content of large video collections, overcoming traditional term-based methods by means of cognitive services such as a Speech-to-Text tool to get video transcripts and the use of Natural Language Processing (NLP) methods to extract semantic concepts and keywords from the above video transcripts [16].

C. Empowering E-Learning Methodologies

Besides, from the methodology perspective, Cen et al. introduce the idea of "big education" applying the paradigm of big data to the whole education process to predict students' performances, based on individuals' learning attitudes and their after-school activities [17]. This seems a promising vein, since also Gudivad theorizes about cognitive analytics driven personalized learning, which can be achieved owing to the advances in cognitive computing for analyzing unstructured data, e.g. blogs, discussions, e-mail, and course messages, to gain insights into student learning at an individual level [18]. New functionalities are strictly connected to new technologies such as, for example, the mix of learning and semantic technologies through the use of ontologies for the description of domain(s) and, in this specific case, the availability of sophisticated cloud infrastructures is required to handle properly such a huge quantity of information, as well as the design and development of new learning environments, supporting suited machine learning technologies [19].

D. Privacy and Security Issues

Finally, we observe that when considering big-data-capable learning applications another paramount item emerges, that is the students' data protection. In fact, personal information in the e-learning frameworks can be very detailed, thus very precise profiling can be obtained and used maliciously for different scopes, such as, e.g., remarketing. As well as privacy, security cannot be neglected and should be considered one of the most important factors in the design of TEL platforms [20]. For example, this topic was faced by Habegger et al., which clearly present possible threats of considering big data within e-learning platforms [21].

III. EMERGING TRENDS IN TECHNOLOGY ENHANCED LEARNING

According to the above considerations, we propose a new point of view on TEL platforms, enlightening the main characteristics they should offer. Their design must consider the new wave of cognitive services for their use in advanced solutions and for their ability to cope with the huge amount of data circulating within the learning frameworks and the connected software environments outside the LMS (e.g., discussion forums and blogs, social network sites, indexing services, digital libraries, etc.). In more detail, personal data about learners, their learning tasks, scores in the assignments, etc. should be stored and historicized to the aim of improving the whole learning and teaching process, also including administrative information to monitor how learning processes are conducted and make an assessment to the aim of predicting performances through suited learning analytics. More precisely, we state that the next-generation TEL platforms will offer pervasive cognitive services.

From the side of the learning process, we have to consider two different aspects: *(i)* learners will acquire new knowledge thanks to the educational strategy and methodology adopted, and *(ii)* at the same time, machines, systems and platforms will acquire information about learners and their individual learning processes, owing to cognitive services. However, thinking of cognitive services merely as the result of an algorithm is not a good starting point: there is something more because learning does not derive from a software algorithm but also from the complex hardware architecture it relies on. In fact, cognitive services can be effective only within a parallel architecture, whose capabilities must be suited to the learning needs of the algorithm itself. In other words, since the algorithm must be trained, we often need substantial memory and computational resources to be implied in educational processes based on a cognitive approach. The personalization of cognitive services may require additional hardware and software resources and thus, for this reason, when defining a cognitive TEL, we have to go beyond a basic hardware configuration to face to possible personalization issues, which may require a system flexible enough to adapt over time its characteristics to the use we will make of it, in the perspective of a lifelong learning support to our knowledge growth. Such features can be achieved through the adoption of cloud architectures, Infrastructure as a Service (IaaS), Platform as a Service (PaaS) and Software as a Service (SaaS) solutions. Furthermore, feedback is needed to improve the effectiveness of this approach, and we should be able to observe the system and get feedback to perform real-time tuning operations.

In the light of the above considerations, it is required to build innovative architectures for the platforms that aim to provide innovative didactics, able to change configuration quickly in order to adapt its functionalities to the different needs of both learners and teachers. In particular, we believe that a wise adoption of the best of some services from multiple vendors, in the same platform, could improve the satisfaction of end users as well as solve some structural problems of the laboratories. In fact, often university laboratories need to be scalable for different number of students attend classes or they may need to serve different educational activities (e.g., quizzes, lessons, exercises, application laboratories, exams, etc.). Sometimes they must even be scaled up to more applicative situations, i.e., performing tests for research projects or experiments. Specifically, Section 4 provides a detailed description of the innovative learning environment that the Federico II University in Naples, Italy, setup, based on the above concepts.

Moreover, we observe that e-learning platforms and applications continuously change together with learning itself and learners' attitude and needs. In many cases, educational paths are designed in collaboration with institutions from other countries, i.e., to promote mobility of both workers and researchers, in other cases a strict interaction with industries and the labor market is required. Consequently, performing educational tasks in a "classroom", whether real of virtual, may be limiting the perspective of what today a smart education should be. Hence, in order to stay competitive, formal education

models should expand their boundaries also involving the external world and this can only be achieved through the adoption of suited technological solutions for online collaboration, interoperability, data exchange, and the seamless integration with legacy systems. The implementation of such solutions requires suited machinery and infrastructures as well as a paradigm shift able to drive the transformation of classrooms into communities to enhance humans' faculties and empowering the transformation of their skills so that, in such new "places", we can speak of men *and* machines, rather than men *or* machines. Even in this case a specific example is presented and Section 5 depicts the characteristics of the innovative e-learning platform developed at the University of Trento, Italy.

IV. An Integrated Environment for Exponential Learning

This section provides a detailed description of the innovative learning environment setup at the Federico II University in Naples, Italy, based on cognitive computing to implement exponential learning. We recall that cognitive computing was born based on human reasoning models and this special capability can be exploited transversally on, potentially, any domain of application, so that it can be regarded as a real revolution, exponentially accelerating processes. In the past decades, we have been accustomed to a swirling growth, according to the Moore's law (the number of transistors in a dense integrated circuit doubles about every two years) or the analogous Metcalfe's law for the number of nodes of a network that clearly explains the enormous dimensions reached by social networking sites such as, e.g., Facebook. However, we were not prepared to cope with the needed exponential growth of learning caused by the growth of knowledge, which is in turn derived by the growth of data processed with novel Artificial Intelligence (AI) techniques. Consequently, the learning environments should be promptly adapted to this reality.

A. The Hardware and Software Setup

According to the above considerations, the Federico II University in Naples, Italy, setup a laboratory with suited hardware and software configurations, providing students of the "Software engineering" class with the possibility to use the most advanced methodologies and tools to develop quite complex projects within their assignments and to share their results with mates. Previously, the authors already carried on didactic activities focusing on computer programming in collaborative environments (see, e.g., [22], [23], and references therein) but cognitive computing has accelerated this need, thus enabling more ambitious projects to be implemented both from the point of view of the hardware resources that can be used and the complexity of the software that can be used. The implemented solution is an integrated environment offering high-level tools for cognitive computing design and programming. Specifically, after a preliminary analysis and a series of stress tests, Microsoft Azure and IBM Cloud have been identified as the best-matching solutions for developing cognitive-computing-based projects. Owing to their cloud-based architecture, both platforms can coexist in the same installation, where they can be used simultaneously if needed, allowing to realize the needed PaaS and SaaS infrastructures.

More precisely, students can use the cognitive e-learning platform from both IBM Cloud and Microsoft Azure with relevant laboratory and customized virtual machines. Taking advantage of such a mixed configuration, enables students to disregard hardware and software issues, focusing on their individual learning tasks. Besides, they can use complex machines simply through a browser, which is a strong point because they may have to use laptops or even personal computers with poor performances. To better clarify, we highlight that the Azure environment offers different profiles for sizing the hardware machine. One can choose one of the following: *(i)* courses, *(ii)* laboratories, and *(iii)* exams. Then, the decision to instantiate a laboratory session or to take part in an exam session is up to the user. The only requirements are the availability of a PC and a network connection. In more detail, the teacher can, in a laboratory consisting solely of the PCs that students bring with them, instantiate an exam session on the fly. Then, the students will have to use their institutional credentials to log into the system, where they can access a personalized dashboard and choose which activity to carry on such as, e.g., laboratory, exams etc. Even if a student is taking an exam, in the same classroom, at the same time, other students can perform different activities. For example, writing and compiling a Java program within the Eclipse IDE as well as entering the IBM Cloud platform, and also use, e.g., the IBM Watson cognitive services.

B. Results

At the end of the activities, the platform allows students to assess the effectiveness and ease-of-use of the environments they used. In particular, they were asked to provide their overall feedback on satisfaction and usability. In order to evaluate the global satisfaction of the platform we considered an ordinal scale with values between 5 and 8. The students were asked to express their opinion after working on the platform for several hours. More precisely, Figure 1 shows that 87% of users gives a medium-high evaluation (scores 7-8), while only 13% believes that the platform has low-sufficient usability (scores 5-6).
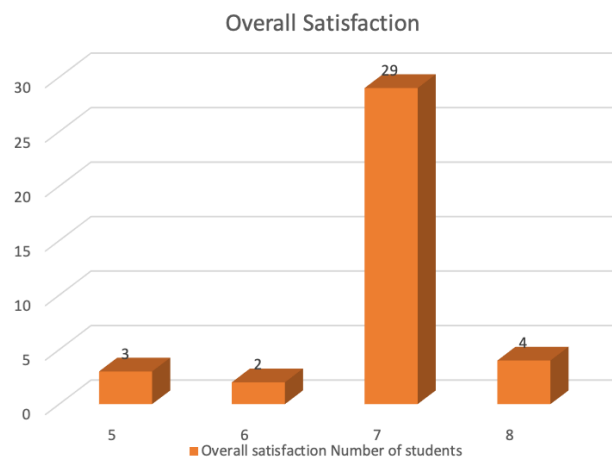

Figure 1: Results from the questionnaires on users' satisfaction.

Since these integrated environments of multiple technologies may be too complex for young students, we have also asked a feedback on usability whose results are shown in Figure 2. In

order to measure the usability of the platform we used an ordinal scale of values between 5 and 9. Students, in a slightly lower number (36 vs. 38), expressed their opinion about the usability of the platform after working on it for the didactic projects they were engaged in. More precisely, Figure 2 shows that the 61% of students gives a medium-high evaluation (scores 7-8-9), while only 39% (scores 5-6) believes that the platform has low-sufficient usability. Besides, students were also asked to express suggestions and possible improvements. From this analysis it emerges that students experienced some difficulties, due to the presence of two different environments but also for the functionalities, rather different from the classical university-laboratory setup they are used to, which offer only basic tools.
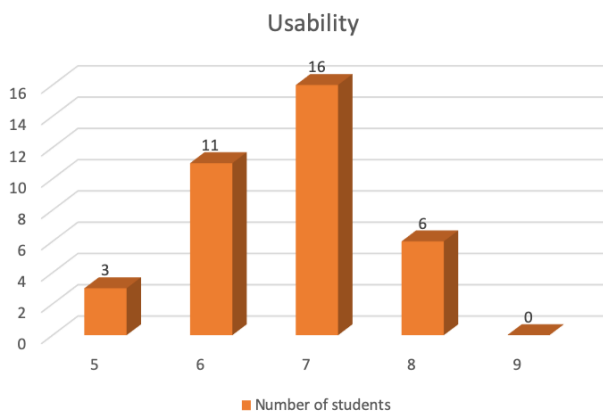

Figure 2: Results from the questionnaires on usability.

In conclusion, the integrated environment in use was judged satisfactory with respect to the functions it had to perform. Given its apparent complexity, we consider such results quite positive and encouraging for the prosecution of this experiment also in the next academic years, in new classes and with a growing number of students.

V. From Classrooms to Communities

According to the above considerations, in this section we depict the characteristics of an innovative e-learning platform developed at the University of Trento, Italy. As highlighted in previous sections, classrooms should shift to communities and in such new places both students and teachers will be able to enhance humans' faculties and empowering their skills transformation, also owing to modern technologies and new paradigms for the possible relationships between men and machines. Specifically, we will discuss the Online Community (OLC) project.

A. The Online Community (OLC) Project

The project started back in 1998 with the idea of implementing a different approach to educational content management, in contrast to proprietary platforms like WebCT™ and Blackboard™, which were dominant at that time. In this scenario, OLC was created from scratch, after having considered three possible alternative solutions: *(i)* adopting a commercial platform; *(ii)* adapting free-open source software to the needs of educational environments, as Moodle or similar where not existing yet *(iii)* building a brand-new platform. Finally, the third

option, despite having the traditional pros and cons of every "make" solution, was chosen for other motivations that now, in the idea of innovating e-learning systems with new cognitive services, revealed to be the right decision. Compared to the adoption of commercial software, the Total Cost of Ownership (TCO) of this kind of solution was probably too high especially in the hypothesis of an extension, not counting the rest of the associated costs for maintenance, management and training, when compared to budget limits. We observed in the following years that this is a very common reason for the adoption of Free and Open Source Software (FOSS) platforms such as, e.g., Moodle by many small-medium educational institutions like high schools and universities. Put simple, being free of charge is the main reason for adopting this solution, without taking care of the side effects on the educational processes that it implies. In fact, a software platform implements processes in a rigid way once they have been coded, so the adoption of this component forces the institution to customize, adapt or even change the way things can be done because "the system does not allow me to do differently". As a secondary effect, the connection of the LMS with the rest of the organization's information system is mostly impossible [24]. As a consequence, many system administrators, are adapting their needs to the software system that, somehow, is able to solve most of their problems, and they mostly are resistant if not reluctant to develop an internal solution. Money, availability of qualified resources, short time to implement the solution, these all are comprehensible reasons for choosing the easy way of acquiring a pre-cooked solution.

We consider much more important to have the possibility of investing in a platform that can be easily extended with new services according to the needs of those trainers willing to experience the use of computer technologies in their educational processes. We therefore decided to develop a completely new platform, and very soon the platform was transformed from a mere LMS to a more structured platform devoted to support collaboration among members of a virtual community. The idea of "classroom" that is lying behind most of the LMS available today is, in our opinion, very restrictive with reference to the more complex processes that normally happen inside educational tasks, and can be extended to any other collaboration environment where collaboration among participants to the community are mediated by technologies. In other words, the idea of a customized LMS that could constitute a competitive advantage for one university versus another one is established [25]. This personalized software is able to supply better and personalized services that ease procedures and processes for the different users such as students, professors, or administrative personnel. Finally, a second decisive item was the need of deep integration with the rest of the legacy information system: authentication with single sign-on, integration with exam records and administrative procedures, possibility of bi-directionally exchange news and messages among people living the university day-by-day routine. Substantially, creating the platform from scratch was related with the rejection of "one-size-fits-all" approach to software components of an information system. Many administrators of the information system, especially in the educational sector, are adapting their needs to the software system that, somehow, is able to solve most of their problems, and they mostly are resistant if not reluctant to develop an internal solution. Money, availability of qualified

resources, short time to implement the solution, these all are comprehensible reasons for choosing the easy way of acquiring a pre-cooked solution.

This adaptation of educational tasks to software platforms is a typical situation where many institutions are lying today. Because teams not always have the knowledge and resources to modify existing (open source) software educational platforms (i.e., LMSs like Moodle), they normally "adapt" themselves to what the platform supplies out-of-the-box, thus limiting the innovation potential of their ideas, and forcing users to adapt their learning processes to the improper technological tool. The typical example is the use of social media (the one that is more appreciated by learners in that moment) to support educational tasks. Social media such as, e.g., Facebook, Twitter, WhatsApp and Instagram are great tools when applied to the context they have been originally created, mostly exchanging multimedia information among peers. Yet, it is not so easy to integrate them in the educational processes: it is clearly a technical problem of available software Application Programming Interfaces (API), but it is also an instructional design problem introducing issues on how to cope the style of the lecture, how is changing the role of the teacher and what are the expectations of learners about the use of social media. In fact, an educational process is something wider than posting a photo or retweeting others' comments, even if it can benefit of this situation. Sometimes, educational processes need the support of other tools and services, that Facebook (for example) can provide through a distorted usage of its services. This normally forces users (mainly educators) to adapt their learning processes to what the platform provides, while it should be exactly the opposite, i.e., the platform should be able to adapt its services to the users' needs. The same happens with LMSs: most of the educational organizations have no possibilities of intervention, nor adaptation or modification on software platforms that have the size and complexity of Moodle, and so they adapt their educational processes to what the chosen platform provides as standard services.

B. Latest Innovations

The innovative aspect that we introduced with OLC, and that now constitutes an extra advantage, is therefore reverting this master-slave approach that sees software platform as masters and end-users as slaves. The approach followed by OLC is to construct from scratch those services identified as relevant by educational experts, based on the precise educational needs of the different users: teachers, students or any other role involved. OLC has some architectural aspects there are very important for our argumentation about a next generation of TEL platforms: *(i)* we own and have created every single line of the source code, so the whole knowledge of the platform architecture and its potential are not scattered among different contributors (like in many open source projects); *(ii)* the platform is equipped with a micro service architecture, thus allowing an easy extension of the platform itself with new parts but taking advantage of the many services that any LMS should provide both to users and to developers that want to extend it; *(iii)* some services have been already developed in the past towards the direction of providing "intelligent" services. Owning and knowing the whole source code of the platform means to have a great advantage if you want to extend it with innovative elements, so the idea of using OLC as a basis for cognitive computing has been straightforward.

Nevertheless, some crucial evolutionary changes had to be applied to the platform, and these changes are the key success factors for this shift:

a) stimulate interaction: the platform should encourage users to interact, not just to download files with teacher's slides. Today, the vast majority of LMSs are used solely to download files, while interactive and more participative services are left to the availability in the platform (if any) and/or at the goodwill of the teacher;

b) pervasive and enriched logging: the platform should log actions of users in order to activate cognitive processes: logging is essential for cognitive computing in order to classify users and infer the best service at the best moment. This logging should not be just the web application server logging, but specialized, application-level logging are needed to capture specific actions inside the single page of the LMS;

c) extensible, service-based architecture: the platform must be extensible through a service-by-design approach, in order to add new services whenever new possibilities can be explored. A micro service architecture is highly recommendable;

d) inference-oriented persistence layer: in order to facilitate inference, reasoning and cognitive computing algorithm, the persistence layer of web platform should be updated to more efficient and flexible data structures [19].

On this basis, a set of profound re-engineering operations has been implemented inside OLC. For example, we extended the platform towards a semantic representation of the knowledge inside the contents of the platform. We also integrated some soft computing, fuzzy-logic-based decision support systems [1], to support decision makers with intelligent tools about educational processes. Moreover, we experimented new storage layers, in order to collect data not only from traditional sources inside the educational environment (i.e., files, forums, galleries, posts, photos, etc.), but also collecting a lot of analytical information about the use of the platform and its services by the users. This immediately opens the problem of the size and appropriateness of traditional relational databases. We performed some experiments in substituting some parts normally stored in relational tables into triples available for a semantic knowledge representation. This meant using triple stores in the beginning as a persistence layer, thus facilitating operations like inference, reasoning, machine learning, etc. The triple (or quadruple) format for persisting (part of) data relevant for decision making and cognitive computing is another step that is not currently available in mainstream LMSs.

To better clarify the above concepts, let us consider as an example the diagram sketched in Figure 3, which depicts the preparation process of proper data sources for big data integration and analysis. The first step is the selection of the data from the persistence of OLC. This persistence is a typical big data source, with structured and unstructured information (file, learning objects, blog posts, forum topics, wikis, etc.) created inside the platform itself. The idea is to separate such data from the rest of the platform, to create the background to be able to apply the cognitive algorithms. So, in a sense, this resembles an Extract Transform and Load (ETL) process, typical in any data warehouse as well as OLAP and data mining solution.
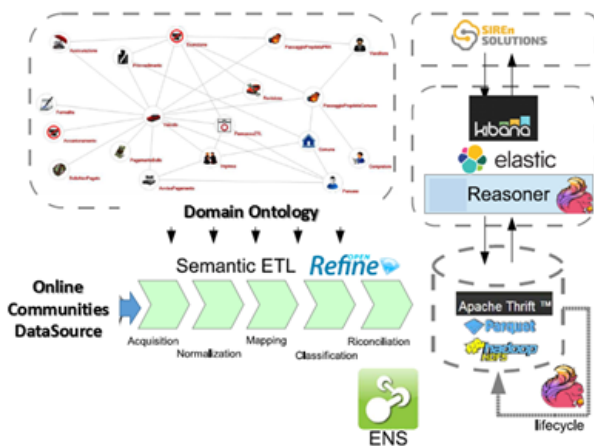
Figure 3: Overview of the big data preparation process for the analytics tools.

The most part of information is coming from the first data source we used in our experimentation, due to its affinity with big data sources, that we call "Actions". This service collects all data coming from users' interactions with other OLC objects or services. In practice, it acts like a sensor introduced inside the source code of the platform in any place the software needs to capture an "action" from the user interface. This is a relevant enrichment of the logs recorded by the web application server, and has been used for many different purposes. Due to volume issues, the system at the moment is blocked on collecting only some types of events, to a certain granularity defined by the system administrator. This choice has not been a design choice, but a performance-related one. In fact, it was clear from the early experimentations that the amount of data could have been compromising the capacity of the DBMS to stand with data acquisition pace and volumes: that is a typical "Velocity" and "Volume" big data problem.

In more detail, there are several elements of data gathering and manipulation pushing our virtual learning environment towards big data, thus increasing the need of a structural change of LMSs architecture, which should adopt new approaches and technologies. Specifically, we mention:

- traditional weblogs, being the application a Web-based software;

- internal logs of usage of the platform, the so-called digital breadcrumbs, that track the learner's journey throughout the entire learning experience;

- mobile logs, where data about mobile learning actions are collected;

- service logs, users' actions on the different elements of the platform like documents, forums, blogs, FAQs etc.;

- logs from the SCORM player, normally an external entity respect to the core services of the platform, with the records of the objects' execution;

- X-API calls, in case the platform is connected or acting as a Learning Record Store (LRS);

- MOOCs, by definition a generator of high volumes of data;

- lifelong learning, an old buzzword of e-learning that is still valid and interesting and, most of all, is another generator of big data, specifically along time;

- serious games that will use materials inside the platform, thus generating a relevant dataset related with users' performances.

The next step in the pipeline is the execution of cognitive tasks which can materialize implicit knowledge to support learning processes. At the moment, this reasoning is limited to basic inference regarding actions performed on certain parts and services of the platform, but the mechanism is ready for larger application scenarios. When the inference process is complete and new knowledge is inferred, a set of administrative routines is executed to load and transform part of the knowledge base to feed applications persistence.

VI. CONCLUSIONS AND FUTURE WORK

As a conclusion, we remark that nowadays we are on a turning point in which cognitive-computing-based approaches are significantly transforming many aspects of our lives. In fact, beside the changes we can notice in professional applications and high-end software and legacy systems, we are already experimenting a set of cognitive computing services in everyday activities, for example, through natural user interfaces and voice assistants, whose presence is becoming pervasive. In practice, more often than we think, we are interacting with machineries that apply sophisticated decision-making processes with very low time constraints and a high level of accuracy. E-learning is one of the fields of application that can mostly benefit from this situation, due to its complexity and to the variety of disciplines that must be adopted concurrently to achieve good learning outcomes. Especially, the use of big data strongly empowers the process of personalization and individualization of the learning processes. Moreover, e-learning is also called to provide suited solutions to the problem of learning to use and exploit such new technologies, which cannot be achieved in environments designed for generic purposes. This raises the problem of developing a new generation of TEL platforms.

The paper introduces the vision of the authors, where a self-made, highly customizable virtual community platform will be integrated with scalable, top-notch cloud platforms and congruent cognitive algorithms applied to the different parts of learning processes, from material selection to educational path suggestions, from peer evaluation to big data discovery for decision makers. The process is still in its infancy, mostly because these three worlds (TEL platforms, cloud services and cognitive computing) are still separated and mostly focus on their own scope. What we are trying to do is merge the three disciplines into one single research area, with precise objectives and deliverables, thus allowing e-learning to maximize the advantages of the fusion of the three.

The two experiences presented in this paper can be regarded as an embryo for the development of future, unpredictable e-learning solutions. To this aim, the results achieved while using the integrated environment for exponential learning deployed at

the Federico II University convinced us to scale the process on a wider and even younger audience, in order to test the real simplicity of such a TEL environment. Another lesson learned is that the surrounding world has become so complex and the changes so rapid that they will never be as slow as in the past. We should not waste our precious time any more in installing software, configuring applications, customizing solutions, tuning databases, maintaining laboratories, etc. Students and teachers definitely have to devote their intellectual power to more creative and less repetitive activities. Moreover, the applications we started using will increasingly have to analyze large amounts of data and therefore be cognitive. Future developments of the OLC also will follow in the direction of providing more advanced cognitive services and making the best from the available information mixing data from different online sources.

REFERENCES

[1] P. Bouquet, and A. Molinari "A new approach to the use of semantic technologies in e-Learning platforms," International Journal of Advanced Corporate Learning, Vol. 9, No. 2, pp. 5–12, 2016.

[2] M. Anshari, Y. Alas, and L. Sei Guan, "Developing online learning resources: Big data, social networks, and cloud computing to support pervasive knowledge," Education and Information Technologies, Vol. 21, pp. 1663–1677, 2016.

[3] M. Coccoli, P. Maresca, and L. Stanganelli, "Cognitive computing in education," Journal of e-Learning and Knowledge Society, Vol. 12, No. 2, pp. 55–69, 2016.

[4] M. Coccoli, P. Maresca, and L. Stanganelli, "The role of big data and cognitive computing in the learning process," Journal of Visual Languages & Computing, Vol. 38, 2017, pp. 97–103.

[5] M. Fedrizzi and A. Molinari, "A multi-expert fuzzy TOPSIS-based model for the evaluation of e-Learning paths." In Proceedings of the 8th Conference of the European Society for Fuzzy Logic and Technology, pp. 554–558, 2013.

[6] L. Banica, and M. Radulescu, "Using big data in the academic environment," Procedia Economics and Finance, Vol. 33, pp. 277–286, 2015.

[7] M. A. Mahmod and A. M. Ali, "Promotion the e-learning success in universities in Baghdad through enhancing their organizational innovative collaboration environment: A qualitative study," International Journal on Perceptive and Cognitive Computing, Vol. 4, No. 1, pp. 12–18, 2018.

[8] B. Dietz-Uhler and J.E. Hurn, "Using learning analytics to predict (and improve) student success: A faculty perspective," Journal of Interactive Online Learning, Vol. 12, No. 1, pp. 17–26, 2013.

[9] S. V. Kolekar, R. M. Pai, and M. M. M. Pai, "Prediction of learner's profile based on learning styles in adaptive e-learning system," International Journal of Emerging Technologies in Learning, Vol. 12, No. 6, pp. 31–51, 2017.

[10] T. Yu and I. H. Jo, "Educational technology approach toward learning analytics: relationship between student online behaviour and learning performance in higher education," in Proceedings of the 4th International Conference on Learning Analytics and Knowledge, 2014, pp. 269–270.

[11] A. Sheshasaayee and S. Malathi, "Impact and consequences of big data in e-learning," in Proceedings of the International Conference on Innovative Mechanisms for Industry Applications, 2017, pp. 726–729.

[12] J. Qiu, Q. Wu, G. Ding, Y. Xu, and S. Feng, "A survey of machine learning for big data processing," URASIP Journal on Advances in Signal Processing, Vol. 67, 2016.

[13] S. Gupta, A. Kumar Kar, A. Baabdullah, and W. A. A. Al-Khowaiter, "Big data with cognitive computing: A review for the future," International Journal of Information Management, Vol. 42, pp. 78–89, 2018.

[14] G.d. S. Leitão, E. B. Valentin, E. H. T.d. Oliveira, and R.d. S. Barreto, "Survey on pedagogical resources recommendation using cognitive computing systems," in Proceedings of the IEEE Frontiers in Education Conference, 2018, pp. 1–7.

[15] D. Dessì, G. Fenu, M. Marras, and D. Reforgiato Recupero, "Leveraging cognitive computing for multi-class classification of e-learning videos," in Blomqvist E., Hose K., Paulheim H., Ławrynowicz A., Ciravegna F., Hartig O. (eds), The Semantic Web: ESWC 2017, Lecture Notes in Computer Science, vol 10577, Springer, Cham, 2017.

[16] D. Dessì, G. Fenu, M. Marras, and D. Reforgiato Recupero, "Bridging learning analytics and cognitive computing for big data classification in micro-learning video collections," Computers in Human Behavior, Vol. 92, pp. 468–477, 2019.

[17] L. Cen, D. Ruta, and J. Ng, "Big education: opportunities for big data analytics," in Proceedings of the IEEE International Conference on Digital Signal Processing, Singapore, 2015, pp. 502–506.

[18] V. N. Gudivad, "Cognitive analytics driven personalized learning," Educational Technology, Vol. 57, No. 1, pp. 23–31, 2017.

[19] P. Bouquet and A. Molinari, "Semantic technologies and e-learning: Towards an entity-centric approach for learning management systems," Journal of E-Learning and Knowledge Society, Vol. 8, No. 2, pp. 65–84, 2012.

[20] L. Caviglione and M. Coccoli, "Smart e-learning systems with big data," International Journal of Electronics and Telecommunications, Vol. 64, No. 4, pp. 445–450, 2018.

[21] B. Habegger, O. Hasan, L. Brunie, N. Bennani, H. Kosch, and E. Damiani, "Personalization vs. privacy in big data analysis," International Journal of Big Data, pp. 25–35, 2014.

[22] M. Coccoli, P. Maresca, L. Stanganelli, and A. Guercio, "An experience of collaboration using a PaaS for the smarter university model," Journal of Visual Languages & Computing, Vol. 31, pp. 275–282, 2015.

[23] P. Maresca, A. Guercio, and L. Stanganelli, "Building wider team cooperation projects from lessons learned in open communities of practice," in Proceedings of the 18th International Conference on Distributed Multimedia Systems, 2012, pp. 144–149.

[24] L. Colazzo, A. Molinari, and N. Villa, "Social networks, virtual communities and learning management systems: Towards an integrated environment," in Proceedings of the 8th IASTED International Conference on Web-Based Education, 2009, pp. 209–215.

[25] L. Kimball, "Managing distance learning: New challenges for faculty," in The Digital University - Building a Learning Community, Springer London, 2002, pp. 27–40.