Journal of Universal Computer Science, vol. 25, no. 10 (2019), 1373-1393
submitted: 15/12/18, accepted: 8/6/19, appeared: 28/10/19 © *J.UCS*

A Web3.0-based Intelligent Learning System Supporting Education in the 21st Century

Khaled Halimi
(Computer Science Department, University 08 Mai 1945 Guelma, Guelma, Algeria
halimi.khaled@univ-guelma.dz)

Hassina Seridi-Bouchelaghem
(Computer Science Department, Badji Mokhtar-Annaba University, Annaba, Algeria
seridi@labged.net)

Abstract: The aim of the paper is to describe the design of a Web 3.0-based Intelligent Learning System (ILS) that addressing the students' needs in the 21^{st} century. The design is based theoretically, on the principles of the connectivism theory and technically, it implements the semantic web representations combining with the use of learning analytics techniques. The work emphasises that implementing a learning analytics approach that uses: text classification, sentiment analysis, topics extraction, and text clustering on the basis of a semantic web and ontologies can support the connectivist learning. The semantic learning analytics process, represents the key element of the proposed intelligent learning analytics system to infer and deduce hidden data in the massive learning data thanks to semantic models of *i*-SoLearn. The aim is to guide students to understand through recommendations, charts and visualisations their learning behaviour and to give teachers feedbacks, enabling them to examine both students' learning and activities. An experimental study using *i*-SoLearn (an intelligent social learning environment), indicates that designing an ILS based on Web 3.0 techniques is effective and expected to show a great advantage in enhancing the connectivist learning of students in the digital age.

Keywords: Computers and education; Semantic Web; Technology Enhanced Learning; Personalization and Profiling; knowledge acquisition; artificial intelligence; Content Analysis and Indexing.
Categories: H.3.1, K.3, L.1.4, L.1.3, L.2.2, M.0, I.2

1 Introduction

The web 3.0 defines a data network that allows computers to understand the meaning of information published on the Web. It extends the network of human-readable web pages to that of machine-readable metadata and creates links between these different contents. This allows intelligent systems to access the web smarter and thus automatically perform tasks in place of users. Therefore, education specialists have found that it is essential to develop intelligent learning systems that take advantage of this shifting, in order to ensure that students of the 21st century will get the best learning experiences according to their needs and preferences.

After the intensive use of Web 3.0 tools, researchers realized that the learning process might occur through users' connections which enable knowledge construction within networks [Kimmerle, 2015]. In this way, new learning theory has emerged: it's the “Connectivism” which aims to understand the learning processes by asking how people use and develop their networks of social relations for their learning and professional development [Downes, 2012].

Since connected learning is a strongly learner-centred and learner-controlled process and with the large amount of data produced by users, learning analytics approaches and the social semantic web technologies present recent trends in analysing the students’ activities which can be very helpful to improve the learning outcomes. These techniques can understand how the learning occurs and develops in networks, how learners create meaning and construct knowledge when connecting with others and how the learning takes place, etc. [Ley, 2016].

In this paper, we propose a new approach that describes how learning analytics combining with the use of semantic web techniques in social learning environment can be applied to develop an intelligent learning environment that supports connected students in the digital age. For this reason, we present a system called i-SoLearn that supports this approach. Therefore, A deep analysis of students’ activities and learning content using the learning analytics approaches, namely: text classification, sentiment analysis, topic extraction and, text clustering, combining with the use of Semantic Social Web techniques is conducted. Since the system's is modelled by a Semantic Web knowledge representation; decision data issued from the proposed analytics approaches are used to infer and deduce knowledge on the learning process. Results of this analytics are provided to students and teachers as recommendations and data visualisation in order to facilitate, support and understand the learning process.

The rest of this paper is organised as follows. In the next section, a brief overview of related concepts of intelligent systems, social learning, connectivism, semantic web and learning analytics is introduced. In section 3, the model of the intelligent system is presented. In section 4, the approach is evaluated through an experimentation. Finally, a conclusion is presented in section 5.

2 Literature review

The web is in constant changing; it has evolved since the earliest days, from the web 1.0 generation that connects people with the Worldwide Web (a network of information connections), going through the web 2.0 generation that connects real people who use the web (a network of people’s connections), to the web 3.0 generation that connects objects of the real people who use the web (a network of knowledge connections) [Rego, 2010]. In 1999, Tim Berners-Lee published the book "Weaving the Web" where he has talked about the future of the web “*I have a dream for the Web in which computers become capable of analysing all the data on the Web the content, links, and transactions between people and computers*". After transforming the paradigms of the Web, from a static to a dynamic Web, an intelligent Web is getting ready, a Web that better understands users’ attitudes and responds exactly as they want. Tim Berner-Lee said the “Semantic web” probably represents the next extreme transformation [Berners-Lee, 2001].

The Web 3.0, which called the semantic web or the web of data is an extension of Web 2.0 combining with the use of recent intelligent computing technologies such as data mining, recommendation agents, machine learning, machine reasoning, big data, natural language processing, cloud computing, linked data, openness, interoperability and smart mobility. This new field of research is central to develop intelligent systems supporting more efficient knowledge management [Vega-Gorgojo, 2015].

The semantic web (SW) shows computers how to understand the meaning of data, this operation will be developed with artificial intelligence agents that can exploit that data. The latter, is defined and linked so that machines can use it not only for display purposes, but for automation, integration and reuse between different computing applications. SW allows software to discharge users while locating relevant resources on the Web, allows web applications to automatically collect resources, to interact with other applications, to perform complex tasks for humans and to access Web resources through content rather than keywords [Passin, 2004].

Education researchers nowadays confidently use the term *E-learning 3.0* in their research [Wheeler, 2009], to indicate the next generation of the intelligent learning systems (ILS) that implement new functionalities of Web 3.0 [Rubens, 2012]. Where an essential shift towards "a more personalized, social, open, dynamic, emergent and knowledge-pull model for learning" is needed to fulfil the new learning requirements of students in the 21st century [Chatti, 2010].

An ILS is a personalized system that places the student at the centre of the learning process, it is based on a seamless integration of shared knowledge according to the semantic web representations and ontologies, the use of rich-media, social and collaboration tools, taking into consideration the students learning preferences, with high use of artificial intelligence, learning data analysis and visualisation and smart-technologies [Gros, 2016]. [Spector, 2014] defined ILS as "*an environment that features the use of innovative technologies and elements that allow greater flexibility, effectiveness, adaptation, engagement, motivation and feedback for the learner*".

Recently, and with the extensive use of web 3.0 and intelligent tools in education, researchers asked if we need a learning theory or new pedagogical approaches to support students of the 21st century? They strongly discussed whether the "connectivism" is adequate as a theory to cover the pedagogical needs in the digital age. With its eight principles, this theory was defined as "*the thesis that knowledge is distributed across a network of connections*". It assumes that being a member of a network, communicating with others and being able to filter information will lead to knowledge creation and progress of learning [Downs, 2012].

A great debate in the literature has emerged, addressing if the connectivism is considered as a new theory of learning. Two tendencies have been developed: the first stream claims that this concept has not offered anything new. For example, [Carreno, 2014] does not consider connectivism as a radical change at the theoretical level, he claimed that it has no big difference with the distributed cognition theory. While the second stream that defended this theory, as [Guité, 2004], for instance, describes the connectivism as "*a model of learning that recognizes the social revolution caused by new technologies*". Despite the significant criticisms and based on the literature review of learning theories, we have examined the important principles of the connectivism and then we have linked the technologies involved under the so-called Web 3.0 to the principles of the theory [see section 3.2]. We argue that the technologies incorporated in e-Learning 3.0 systems are effectively corresponding with the principles of the connectivism. Therefore, we join the second stream since we strongly believe that the connectivism could be the best learning theory that supports learning in the 21st century.

In a web 3.0-based learning environment, anyone can anytime participate in the creation and sharing of knowledge, regarding the massive amount of data generated by students through their activities and interactions, it is challenging to support and assist all of them during their learning. Consequently, it is essential to rely on artificial intelligence and data mining technologies in order to check, sort and analyse the educational data, etc. [Rubens, 2012]. Researchers add the concept "anyhow" to the concept of "anytime, anywhere and anyone" to leverage the intelligence that characterizes the E-Learning 3.0 systems. This "anyhow" feature is highlighted by the new research field termed "Learning Analytics" [Rego, 2010].

The term Learning Analytics commonly refers to computer, mathematical, and statistical techniques to extract relevant information from very large sets of data. [Siemens, 2012] views LA as "*the use of intelligent data, learner-produced data, and analysis models to discover information and social connections, and to predict and advise on learning*". The objective of this domain is to monitor and analyse all data produced by learners, with the aim of personalisation and adaptation of learning with a predictive dimension [McAfee, 2012]. In general, the analytics is used to give meaning to the activity of the learner [D'mello, 2012]. Learning Analytics combines concepts used in web analysis and data mining, among the used methods we can note:

- *Semantic Representation*: The main idea behind the Semantic web seems very appropriate to meet the new needs of learning analytics, the fact that the analysis process leverages the meaning expressed by semantics presented in relationships between concepts of the ontology and the discovery of new knowledge by means of the inferring mechanisms. Ontology improves the analysing accuracy when searching for resources, avoiding terminological ambiguities [Halimi, 2014].
- *Predictive analysis*: enables to make assumptions about future events based on the observation of past facts. The main hypothesis is that the capture of relations between explanatory variables and explained variables, will make it possible to predict the future explained variables [Fancsali, 2011].
- *Content recommendation*: considers not only the similarity of items sought, but also the similarity of the consultation history of students, it can identify relevant contents to highlight, and individualize the selection of content proposed to a particular student [Khribi, 2008].
- *Community detection*: relies notably on social networks analysis, allowing to identify homogeneous clusters in the network, to investigate the density of these networks and to identify the egocentric network, etc. [Fortunato 2010].
- *Text-mining*: involves extracting new knowledge from weakly structured data (log files, conversations, posts, messages, etc.). The automated analysis of the members' contributions requires techniques of natural language processing, namely: topics classifications, text summarization, sentiment analysis, opinion analysis and discourse analysis, etc. [Montalvo, 2018].
- *Machine-learning*: it is about designing learning systems capable of reasoning quickly on weakly structured data. The idea consists of developing algorithms simulating the functioning of human reasoning: Bayesian inferences, neural networks, memorization, case-based reasoning, etc. [Brusilovsky, 2003].
- *Data-visualization*: is an integral part of a data analysis process that provides users with the ability to explore numbers, providing keys for critical reading, and the opportunity to discover unexpected elements. Dashboards provide a visual

interpretation of large data sets to discover, interrogate, understand the elements of information that are difficult to understand in a written text [Verbert, 2013].

Some research works have been carried out leveraging the use of Web 3.0 techniques in order to enhance the design of intelligent learning system. [Ferguson, 2015] evaluated the clustering method to determine engagement models in a MOOC. [Aguiar, 2014] proposed a method based on classification algorithms to measure student engagement from their digital portfolios to evaluate the quality of prediction. [Martin, 2013] established different paths in learning fractions with online gambling, based on the use of data mining methods, graph visualizations and a classification algorithm. [Joksimovic, 2015] measured the influence of discourse models on the social capital of learners. They used non-directed graphs and linguistic analysis to measure the quality of writing of learners in a cMooc. [Snow, 2015] proposed a methodology based on NLP techniques (narrativity and cohesion) to evaluate the stylistic flexibility of confirmed writers. [Martinezmaldonado, 2015] developed an iterative process of designing, validating and deploying visualization tools to focus the teacher's attention on the group collaboration and progression. MeLOD[1] is a system that supports analytics of learners' activities in a mobile learning setting based on the Semantic Web. PBL3.0[2] integrates learning analytics and semantics in problem-based learning. AFEL[3] provides analysis and understanding of learner's data to personalize and enhance the learning process. Didactalia.net[4] is a storage space for teachers, students and parents to create, share and find open learning resources with semantic contexts, providing analytics and data visualization. WATCHME[5] uses learning analytics to improve workplace-based feedback and professional development with data visualizations. PredictED[6] analyses student behaviours in Moodle. CourseSignals[7] is predictive learning analytics which uses student data to predict those who are at risk. CognitiveTutor[8] is an intelligent tutoring system implementing learning analytics and reporting. FFTAspire[9] is a data analysis and reporting tool providing many dashboards showing facets of school performance.

We conducted a comparative study between approaches and systems mentioned above taking into consideration the different services and features they provide. Table 1 summarises the most important characteristics of the studied systems. Therefore, based on this comparison, concepts and theories discussed earlier, we found that there is a lack of the research on relating semantic web to learning analytics approaches, which represents a key element for the development of the future intelligent learning systems. Therefore, our contribution in this work is articulated around an original idea of analysing the students' textual contributions through Semantic web modelling, and to present a model of an intelligent learning system (*i*-SoLearn) that implements Web

[1] MeLOD: http://melod.pa.itd.cnr.it/
[2] PBL3.0: http://hdl.handle.net/1820/8701
[3] AFEL: https://projects.kmi.open.ac.uk/afel/
[4] Didactalia.net: https://didactalia.net/
[5] WATCHME: http://www.project-watchme.eu/
[6] PredictED: http://cloudworks.ac.uk/
[7] CourseSignal: https://purduestudio.org/
[8] CognitiveTutor: www.carnegielearning.com/
[9] FFTAspire: https://fftaspire.org/

3.0 technologies in order to improve and enhance the connected learning in the digital age. A detailed description of the *i*-SoLearn is presented in the next section.

Systems / Approaches — Characteristics	WatchMe	PredictED	Course Signals	Cognitive Tutor	AFEL	itslearning	MeLOD	PBL3.0	Didactalia.net	FFTAspire	Joksimovic, 2015	Aguiar, 2014	Martin, 2013	Kizilcek, 2013	Charleer, 2014
Web 2.0 tools							√		√		√			√	
Semantic Web							√	√	√						
Recommendation						√	√							√	
Feedback	√		√	√	√				√	√	√				
Personalization				√										√	
Inferring				√											
Learning Analytics	√	√	√	√	√	√	√	√	√	√	√		√		
Visualization	√			√	√	√	√	√	√	√			√		√
Own data	√	√	√	√	√	√	√	√	√	√			√	√	√
External data						√					√	√			
Prediction		√	√	√			√					√			
Classification		√									√	√	√		
Self-regulation		√													
clustering													√	√	

Table 1: Comparison between the main characteristics of Web3.0-based systems.

3 Model of the Web 3.0-based intelligent learning system

Having analysed the state of the art in domains of Web 3.0, as well as the state of the practice in the application of Semantic Web and learning analytics in education, a description of a *i*-SoLearn (a Web 3.0-based Intelligent Learning System) which is based on the principles of the connectivism theory and implements the semantic web representations combining with the use of learning analytics techniques is presented in next sections.

3.1 System overview

i-SoLearn is an intelligent learning environment that supports students in the digital age. Firstly, an overview of its theoretical foundation is presented. Then, the knowledge representation model, the Web 3.0-based pedagogical model, the applied learning analytics approaches and a real pedagogical scenario are also detailed.

Generally speaking, the proposed system aims to establish an online meeting space between teachers, experts and students; it enables users to create groups of interests with different levels of knowledge and personalisation. Within its connectivist perspective, the learning entails greater participation of students, allows resource sharing, such as videos, photos, and links, permits communication via mail, chat or video, allows users to ask questions, give answers, share posts, and receive notifications, feedback and recommendations. Within its data analysis and visualisation tools; teachers, learning staff, and students can easily monitor the learning activities and extract interesting conclusions that can enhance the learning experiences.

3.2 The Theoretical Background of the system

The theoretical foundation of the system is based on the principles of the connectivism theory [Simens, 2012]. We present in the following the application of these principles according to our vision:

1. Learning and knowledge rest in diversity of opinions: the system is developed as a social learning network enabling students to interact through discourses, debates and critical dialogue using web 2.0 tools.
2. Learning is a process of connecting specialised nodes or information sources: the system is a set of networked nodes that represent users, courses, strategies, objectives, etc. Learning will take place when connections between entities occur, for example, when a student asks teacher or tutor for help.
3. Learning may reside in non-human appliances: the knowledge can be found in learning networks, learning objects, tags, wikis, links, opinions and relationships between all these sources supported by the ontology.
4. Capacity to know more is more critical than what is currently known: easy access to knowledge is guaranteed by the learning services, which allow students to have much control over knowledge and how to apprehend it.
5. Nurturing and maintaining connections is needed to facilitate continual learning: through the intelligent recommendations provided by the system, the student never strays from his/her objectives as the system encourages him/her to interact, communicate and participate in the knowledge creation and sharing.
6. Ability to see connections between fields, ideas, and concepts is a core skill: a concept-map is used to facilitate the visualisation of the semantic representation of concepts and their relations presented in the ontology.
7. Currency is the intent of all connectivist learning activities: with a set of semantic inferring rules, the system provides appropriate knowledge that fits well with students' needs according to their activities and styles.
8. Decision-making is itself a learning process: the system provides students with the personalisation of their learning paths and content over a free decision to choose what, when, where, with who and in what way they will learn.

3.3 The conceptual model

In the following, the system's conceptual model and all its components are presented in detail. As shown in figure 1, all the knowledge is modelled by Semantic Web representations, then a deep analysis of students' activities and their content using analysis tools combining with the use of Semantic Social Web techniques is performed. The results of this analytics will be provided to students and teachers as recommendations and data visualisation to support, understand, and draw conclusions about the learning experiences.

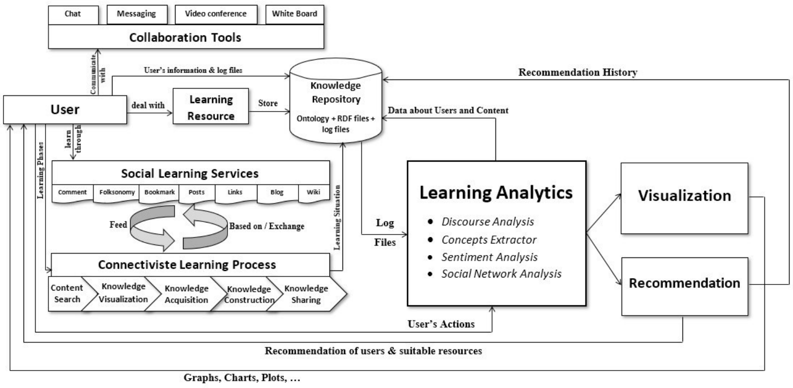


Figure 1: The conceptual model of i-SoLearn.

3.3.1 Knowledge representation

We have used an ontology already developed in a previous project [Halimi, 2015]. In this work, we have modified it to support the semantic annotations of the connectivist learning objects, learning strategies, knowledge creation and the different analysis processes. Figure 2 shows the ontology that describes three types of knowledge: concepts, properties, and individuals; it is composed of different classes and sub-classes. Part A describes knowledge about users, learning objects, learning styles, learning domain, and learning objectives, etc, part B describes properties defining the objects and part C describes the Individuals or the instances of the objects.

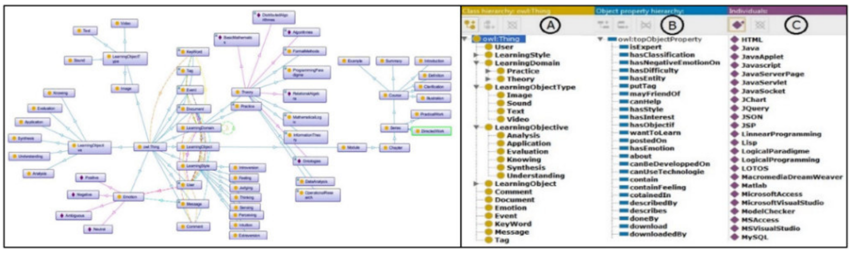


Figure 2: Extract of the ontology.

3.3.1..1 The Student's model

The student's model is an RDF file containing: personal information, cognitive level, mastery of learning domains, learning style, social networks, messages, tags, and sentiments, etc. The User class is related to other classes through a set of properties. For example, to express the student's learning objectives, the property *hasObjectif* is used. The Student class is related to *Learningdomain* class and *LearningStyle* class through the *WantToLearn* and *hasStyle* properties. Emotions of students are presented using the property *hasEmotion*, etc. Figure 3 shows an extract of the student RDF model using the N3 syntax.

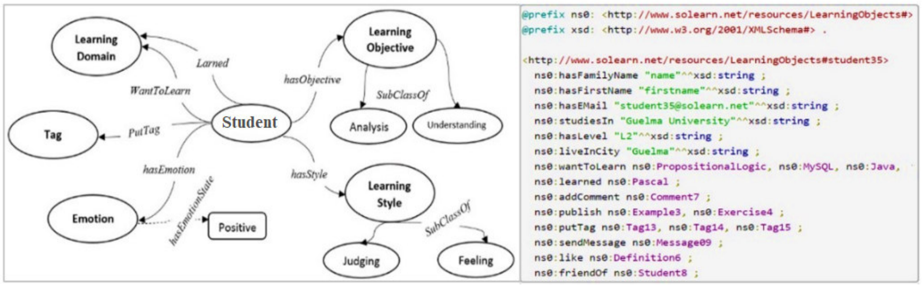


Figure 3: The student model with serialisation in N3 syntax.

3.3.2 Web 3.0-based pedagogical model

To support modern learning methods that match the needs of students in the digital age, by which everyone gets at home, at school or outside the correct and desired answers just in time and as quickly as possible, through connections with different sources of knowledge. We designed the pedagogical model as a connectivist form with a semantic representation to support sharing, reusability, and flexibility. The model is based on dividing learning resources into a set of learning units, each unit is divided into a set of modules, and each module is also divided into a set of courses; each course contains a set of learning entities (definition, summary, illustration, example, etc.), each entity could have different formats (for instance, definition is a text file while illustration is a video). We have defined also semantics relationships between these entities, for instance <definition, *isPrerequisiteOf*, illustration>. This design takes into account the students' differences, where a single course structure cannot be appropriate for all students. This decomposition approach is advantageous in the sense it allows for quick retrieval of knowledge and does not force students to consume a content that does not fit with their preferences.

3.3.3 Web 3.0-Based Connectivist Learning Process

We proposed a connectivist learning process that has a robust relationship with social learning services [see Fig. 1]. On the one hand, the latter feed the learning process with the necessary knowledge created and shared by "*expert users*". On the other hand, it's the learning process that stimulates users to use social learning services, e.g., it's the "*content search*" module that prompts students to use the folksonomy in order to find content or users in relation with certain learning domain, or it's the "*knowledge creation*" phase that drives students to create and share learning objects through wikis or blogs, etc. With the semantic "*Content Search*" module students get a "*big picture*" of their learning objectives. The model extracts all knowledge stored in the ontology as concepts and relationships between them so that students will be able to find where and how to learn. By using a concept map for "*Knowledge visualisation*", students who do not have any knowledge about a specific topic can rapidly get its general meaning and relate the new and old information. For "*Knowledge Acquisition*" students can at any time get appropriate learning resources or get answers and feedbacks in the form of posts, comments or messages from the "*more knowledgeable others*". For "*Knowledge construction*" users can easily create knowledge by using the social services to communicate or discuss any issue related to their learning through which ideas, concepts or practices are shared in networks.

All knowledge about content and users' actions generated in all previous phases will be stored in the knowledge base as RDF files, to use them by the learning analytics module, which will be presented in the following section.

3.4 Learning Analytics

In this section, a description on how the analytics approaches are applied to support the learning, and how they get the benefit of the semantic web models is presented.

Semantic learning analytics can be used for viewing and analysing massive data generated by students, which helps them to understand their learning and behaviours. It also gives teachers means to detect where students have weaknesses or limitations and where their courses are inappropriate. Tracking students' activities using semantic models gives the analytics processes the advantage of exploiting the meaning expressed by semantics contained in relationships between concepts defined in the ontology, and enabling further inferences on the data. Thus, the semantic analysis supports the learning requirements of the intelligent system, by answering questions like: Who works with whom? When do they work? What did they do in the past? What resources do they access? What do they know? What do they prefer? What will they do in the future? How do they feel? Etc.

The analytics process is based on the analysis of the students' RDF files (which contain all the actions and activities carried out during their learning (such as: adding resources, commenting, sending messages, performing tests, etc.). Thanks to the inferring and reasoning mechanisms of the semantic web, new knowledge and facts about the students' activities will be extracted from their graphs, providing the opportunity to discover unexpected elements that can be useful to improve learning. We present in table 2 the learning analytics approaches applied in the system.

3.5 An application scenario to analyse learning content

In the following, we'll see how both teachers and students get the benefit of the analysis process through the following scenario. Let's suppose that a teacher basing on the pedagogical model presented before [see section 3.2.2], adds the course1 that comprises a set of pedagogical entities: introduction, example1, example2, definition, exercise1, exercise2, and exercise3. As shown in figure 4, example1 has the format Video, defined semantically with the triplet: *<lo:Example1 lo:hasFormat lo:Video>* and example2 has the format Text.

Students during their learning perform several actions on course1, they comment, download, reuse, tag, share, like, dislike any of its entities. Analysing these actions through the learning analytics module will give conclusions to both teachers and students as presented in the following.

Objective	Data	Technique	Results
Social Network Analysis Identify interactions which promote the learning process.	Student's RDF files	**Social Network Analysis** (*Centrality, Betweenness, etc.*.)	• Identify students who get the most friendship requests; • Who tag a resource already tagged; • Who worked on same domain; • Who have same learning style; • Students who influence their networks more than others; • Students have positive or negative reactions.
Sentiment Analysis Extract the sentiments which describe the emotional state of students towards learning objects or other users' contributions	Students' texts stored in their RDF files (posts, comments, messages, etc.)	**AYLIEN Sentiment Analysis API** *(an intelligent tool that uses Machine Learning and NLP) Followed by* **SPARQL Query Ontology**	• Decide whether students' texts have a positive, negative or neutral emotion; • Text is subjective or objective; • Information about the students' impression about teachers' courses; • Detect which part of the course has a positive or negative reaction.
Discourse Analysis Discover what students knew, understood or misunderstood.	Students' texts stored in their RDF files (posts, comments, messages, etc.)	**uClassify Discourse analysis API** (*a free machine learning web service enables creating and using text classifiers.) Followed by* **SPARQL Query Ontology**	• How learners organise the acquired knowledge; • Check if they understand the new terminology after it has been used in the course; • Verify whether they are including the new terminology as part of their vocabulary; • Review frequently used words. • Identify why students' discussion did not intersect with the content of the course; • Find where students give critics, judgments, right answers or ask questions, etc.
Learning resources analysis Understand the main content of resources shared by students	Students' texts stored in their RDF files (posts, comments, messages, etc.)	**AYLIEN Concept Extractor API** (*an intelligent tool that uses Machine Learning and NLP*) Followed by **SPARQL Query Ontology**	• Understand the main content of resources shared by students; • Determine the learning domain of resources; • Determine the learning objective of content; • Determine learning strategies • Determine relations with other objects; • Determine knowledge type of the shared content: Know what, know why or know how.

Table 2: The applied learning analytics approaches in the system.

3.6 An application scenario to analyse learning content

In the following, we'll see how both teachers and students get the benefit of the analysis process through the following scenario. Let's suppose that a teacher basing on the pedagogical model presented before [see section 3.2.2], adds the course1 that comprises a set of pedagogical entities: introduction, example1, example2, definition, exercise1, exercise2, and exercise3. As shown in figure 4, example1 has the format Video, defined semantically with the triplet: *<lo:Example1 lo:hasFormat lo:Video>* and example2 has the format Text.

Students during their learning perform several actions on course1, they comment, download, reuse, tag, share, like, dislike any of its entities. Analysing these actions through the learning analytics module will give conclusions to both teachers and students as presented in the following.

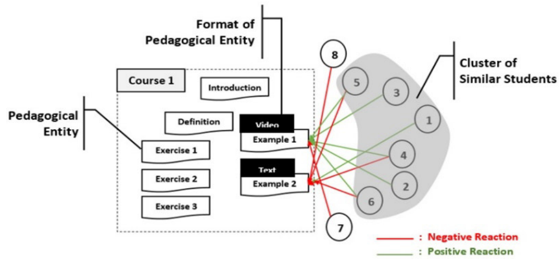


Figure 4: Illustration scenario.

Based on the principle of the proposed Web 3.0 pedagogical model, any student during her/his learning can at any time share any learning resource within his/her community. Therefore, understanding the main content of resources shared by students is extremely necessary. With the use of the AYLIEN Concept Extractor API, as explained in table 2 and as shown in figure 5, the system extracts the concepts from any shared text (learning object, posts, answers, etc.), then subjected them to the ontology, through the use of the semantic relations, the inference rules and SPARQL queries, the system can understand the content of resources (learning domain, objective, relation with other resources, etc).

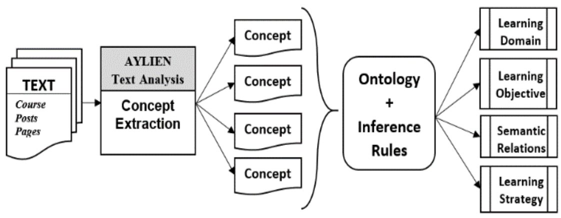
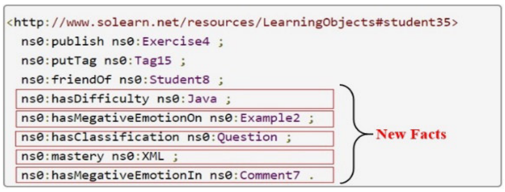


Figure 5: Learning Resource Analysis

On the one hand, this analysis is used to discover what the students' content is exactly about. On the other hand, teachers can use this module to evaluate the students' contributions. For instance, a teacher asks students to upload pedagogical entities about "Java". The system classifies the students' pedagogical entities and subjects them to the ontology to extract all the knowledge in relation with them. Suppose that the system detects that the *summary7* of *student4* talks about "Java Applet" which has a "required" semantic relation with the object "Java" as defined in the ontology: <lo:JavaApplet lo:hasRequirement lo:JAVA>. There is a semantic relation, the system concludes that *student4* is in the right direction to understand the course. Suppose now that *student11* added *example9* that talks about "JavaScript", the latter is described in the ontology as:<lo:JavaScript lo:isDifferentFrom lo:JAVA>, the system concludes that *student11* is completely away from the course's objectives.

Based on results of this analysis, the system automatically sends to teachers a list of students who have understand the course objective and a list of students who faced difficulties in assimilating the goals of the learning content and recommend to students the appropriate learning resources that better fit with their needs.

Figure 6 shows the new knowledge generated by the previous analysis processes, this knowledge will be stored in the students' RDF files as new facts to use them by the intelligent recommendation module.



```
<http://www.solearn.net/resources/LearningObjects#student35>
  ns0:publish ns0:Exercise4 ;
  ns0:putTag ns0:Tag15 ;
  ns0:friendOf ns0:Student8 ;
  ns0:hasDifficulty ns0:Java ;
  ns0:hasMegativeEmotionOn ns0:Example2 ;
  ns0:hasClassification ns0:Question ;
  ns0:mastery ns0:XML ;
  ns0:hasMegativeEmotionIn ns0:Comment7 .
```


Figure 6: Inferring new facts.

3.7 Intelligent Recommendation

The intelligent recommendation is one of the most important and innovative features of *i*-SoLearn. The recommendation module which is based on results of the analytics approaches presented above combining with the use of inferring and reasoning mechanisms of the semantic web allows us to decide what type of learning resources or users must be proposed to students to help them achieve their learning goals. In the field of Artificial Intelligence, inferring is a component of the system that applies logical rules to the knowledge base to deduce new information. In the following, some inferring rules based on SWRL (Semantic Web Reasoning Language) are presented:

- *isExpert (?x, ?d)* ∧ *wantToLearn (?y, ?d)* → *canHelp (?x, ?y)*

If student *x* is an expert of learning domain *d* and if student *y* wants to learn the same domain *d*, then the system infers that user *x* can give help to learner *y*. Therefore, the system will recommend to *y* to add user *x* to his/her social learning network.

- *putTag(?x, ?t)* ∧ *putTag(?y, ?t)* → *mayFriendOf (?x, ?y)*

If two learners have used the same tag on the same learning object, then these two learners can be friends. Therefore, the system recommends them to add each other.

- *hasStyle (?x, ?sensory)* ∧ *WantToLearn (?x, ?d)* → *hasEntity (?d, ?example)*

If student *x* has the sensory style, and s/he wants to learn domain *d*, then the system recommends to *x Examples* about domain *d*.

- *post (?x, ?c) ∧ postedOn (?c, ?d) ∧ hasClassification (?c, ?question)→ hasDifficulty (?x, ?d)*

If student *x* posted a comment on document *d* and after subjecting the comment *c* to discourse analysis module, the system discovers that *x* asks a lot of questions about *d.* Therefore, it infers that *x* has difficulties about *d* and will recommend to *x* all users mastering *d* and all resources in relation with it.

- *post(?x,?p)∧postedOn(?p,?lo)∧(isExpert(?y,?lo)∧ hasNegativeEmotionOn (?y, ?p)) → hasLowLevel (?x,?lo)*

If student *x* added a post on a learning object *lo* and user *y* which is an expert in the domain of *lo* and has a negative emotion on the post added by *x*, the system could infer that *x* has a low level on that domain.

- *isNovice(?x,?lo)∧hasNegativeEmotionOn(x,?lo)) → hasDifficulty (?x, ?lo)*

if *x* is a novice in a learning domain *lo* and s/he usually shows negative reactions about resources treating *lo*, the system infers that *x* has probably difficulties about *lo.*

- *putTag(x, t) & putTag(y, s) & sameAs(t,s) → mayFriendOf(x, y)*

If a user put a tag on a learning object and another user put a different tag on the same object, but the second tag has a relation like *sameAs* or *hasRequirement* with the first tag, each user is recommended to the other.

3.8 The system's Dashboard / data visualization

The *i*-SoLearn Dashboard that provides a set of visualisation tools was designed to help students, teachers or learning staff to monitor learning activities such as social activities, learning outcome, learning progress, etc. and to extract interesting conclusions that can be used to enhance learning experiences.

Figure 7 shows charts available to users: a bar chart (A) represents the discourse analysis results. It shows how much a student gives questions or disagreements during his/her learning. This information can help students to be more motivated and look forward to finding ways to be more positive. A vertical bar chart (B) represents the student's social actions: commenting, messaging, tagging and friendship request. A poll bar chart (C) represents the student's dominant learning style. The gauge chart (D) shows the student's social state at a given period. A spider chart (E) shows the covered learning domains used in the community, thanks to the inference possibilities of the semantic web, the system with a generalisation relation and simple SPARQL queries can automatically infer learning domains that interest students. Teachers can identify which students are not participating actively in the community and get clues to support analysis of factors underlying this lack of active participation. For example, a student's relationships with his/her peers to evaluate the risk of isolation; a student's difficulties while using new web 2.0 tools. Type of activities undertaken by each student can also provide information his/her preferred learning style.

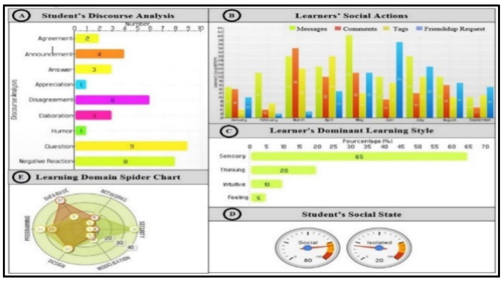


Figure 7: The i-SoLearn Dashboard.

4 Experiment

An experimental study is in progress using a prototype of *i*-SoLearn (a system under development). The purpose of the experimentation which is divided into two parts is to demonstrate the effectiveness of using Web 3.0 tools to enhance the students' learning. For this end, we used a Simple Random Sampling where we have chosen randomly 26 students from the entire active students subscribed in the system. Topic taken was general concepts on computing. Feedbacks of system's usage was obtained through a survey (we used a questionnaire composed of 20 questions to gather data about individuals), the samples were subjected to two different experiments.

In experimentation 1 (Before-After situation): we tried to verify if the Web 3.0 technologies (especially, semantic web and learning analytics') have a positive effect on the connectivist learning process. We conducted the experiment on two phases; the first one was done using a simple learning environment providing only a space to exchange resources and discussions. While the second phase was done using the system with all its features: social networking, semantic search, inferring, learning analytics, data visualisation and recommendations, etc.

In experimentation 2: we tried to check whether the student's sentiments towards a learning entity determine its level of difficulty or its inappropriate design, where we have used the AYLIEN Sentiment Analysis API to analyse students' comments.

4.1 Methodology

4.1.1 First experimentation

In order to know the effectiveness of the approach, we proposed to test the following hypothesis: Null hypothesis H0: *the use of Web 3.0 techniques does not have any effect on enhancing the connectivist learning process*.

To verify this hypothesis and after aggregating all students' data (number of posts, tags, friendship invitations, collaboration, contact search, comments, etc.) before and after using the proposed approach, we compared the averages of the two paired values where both observations were taken from same participants. To determine if the difference between the two averages is greater than zero, we used the t-test as the sample size is less than 30.

4.1.1. Results of the first experimentation

After using the Excel Analysis Pak which is a free Excel component, we obtained results presented in table 3 with a confidence level of 95%:

Situation	N	Mean	Standard deviation	T_{score}	P_{value}
Before	26	4,072	2,878	2,711	0,0010
After	26	7,557			

Table 3: Results of the t-test experimentation

According to results on Table 3 $t_{0.95} = \pm$ 1,162, the mean (M=3,485, SD=2,878, N=26) was significantly greater than zero. t(26)= 2,711, P_{value}=0,0010 and a 95% C.I. about Mean is (2,323, 4,648), providing evidence that null hypothesis *H0* is rejected in favour of the alternative hypothesis. Therefore, the Web 3.0 techniques are effective in enhancing the connectivist learning process.
In Figure 8, students' responses to some questions in the questionnaire are presented.

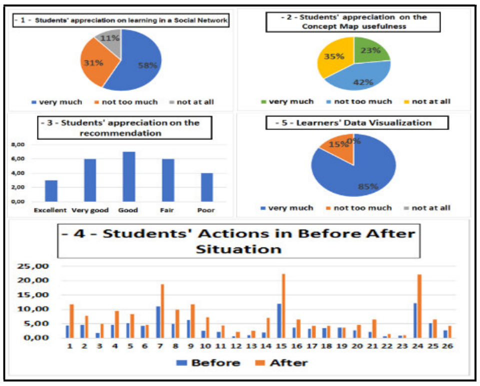


Figure 8: Students' responses to the survey questions

4.1.2 Second experimentation

In this second experiment, we have focused on the verification of the question: "*Is there any relation between the difficulty or the design of a pedagogical entity and the student's sentiments expressed in comments about that entity*?" We asked students to download an exercise proposed by a teacher on the system, it is about "Java classes" and which is a bit difficult compared to their levels, then we asked them to add Comments to their spaces evaluating the difficulty of the exercise, we subjected the students' comments to the sentiment analysis module in order to extract their emotion towards the exercise. At the same time, we asked them to answer this question on the printed questionnaire: "*How do you evaluate the difficulty of the exercise*?" Where students' answers vary from "very easy, easy, neither easy nor difficult to very difficult", with corresponding values (-2, -1, 0, +1 and +2).

4.1.2. Results of the second experimentation

Results of the sentiment analysis process and results of the printed questionnaire are presented in Table 4.

Pedagogical Entity		**Course: Java || Entity: Exercise3**	
Evaluation of the exercise by students			
N^br	**Results on the system**	**N^br**	**Results on the questionnaire**
1	P+: strong positive	**0**	Very Easy
3	P: positive	**3**	Easy
6	NEU: neutral	**3**	Neither easy nor difficult
9	N: negative	**11**	Difficult
7	N+: strong negative	**9**	Very difficult

Table 4: Results of the test on the course's difficulty.

Figure 9 shows the comparison between the results of the sentiment analysis and the student responses to the questionnaire.

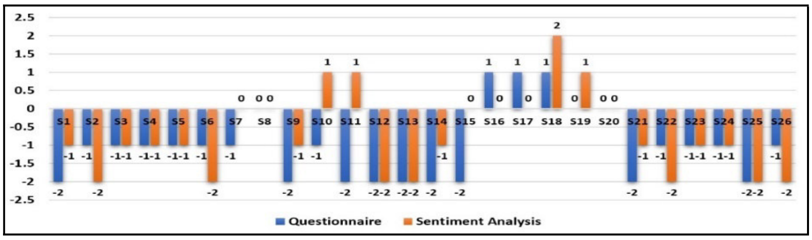


Figure 9: Results of the sentiment analysis vs. responses to the questionnaire.

4.2 Discussion

Throughout the first experiment, on the one hand, we have established that the use of different Web 3.0 technologies has prompted students to participate much more. Therefore, we argue that the learning will be enhanced based on the principles of the connectivism's philosophy. Moreover, learners have appreciated so much the prototype of the system, especially its design as a social network. The most important feature that impressed everyone is the results of the different learning analytics approaches. A student, for instance, said about the result of his discourse analysis: "*I was really surprised how much I gave critics…*" another one argued about her Sentiments: "*I did not imagine I'm negative to this point*", another learner (23), by using the social state gauge finds that he is very "isolated". As we can see in Figure 8, the student 24 almost doubled his activities in the system (Rise by 83%) after showing him his statistics. On the other hand, we admit that students who have little interest in social networks and web 3.0 tools have not really got the benefit of the system's possibilities (students 6, 19, 22, 23). To a large extent, we recognise that this approach requires students who are very proficient in using modern technologies to achieve their learning objectives.

Throughout the second experiment, results of the proposed approach show that, to some extent, there is a significant convergence between the results of the sentiment analysis and the students' answers on the questionnaire [see Fig. 9]:

- *On the questionnaire*: 12% said that the exercise is easy; 12% said that it is neither easy not difficult and 76% said that it is difficult (where 34% said it's "very difficult" and 42% said it' "difficult").
- *With the sentiment analysis module*: 16% have a positive sentiment (where 04% have a P+: "strong positive sentiment" and 12% have a "P: positive sentiment"); 23% have a neutral sentiment and 61% have a negative sentiment (where 34% have an "N:negative sentiment" and 27% have an "N+:strong negative sentiment").

In general, the following conclusions can be drawn:
61% of students (16) have evaluated the exercise as difficult on the questionnaire and they showed a negative sentiment about the exercise on the system; 32% of students (08) have evaluated very precisely the exercise on the questionnaire as on the system (03 students: [N+ Strong Negative {system} & Very Difficult {paper}] and 05 students: [N Negative {system} & Difficult {paper}]) and 07% of students (02) have evaluated the exercise as difficult on the questionnaire, but they showed a positive sentiment on the system.

By deleting students who showed a "Neutral Sentiment" (06 students), the rate of students who evaluated the exercise as "difficult" on the paper and on the same time showed a negative sentiment towards it on the system has increased from 61% to 80%. Therefore, we conclude that the student's sentiments towards a pedagogical entity clearly determine its level of difficulty or its inappropriate design.

Figure 10 shows a mock-up of *i*-SoLearn (currently under development), which recommends personalised learning resources and provides a lot of visual statistics of students' activities based on a set of learning analytics approaches.

Figure 10: Some interfaces of i-SoLearn.

5 Conclusion

In this paper, we presented how Web 3.0 technologies can contribute to the development of intelligent systems that in turn can play a key role in improving the quality of the connected learning in the 21st century. Basing on the use of semantic analytics approaches, learners, teachers and learning staff can understand the data generated through the various learning activities and can get answers and proper tools to enhance the learning outcomes. We showed how the learning analytics approach that uses: text mining techniques can rely on the social semantic web models, the fact that the analysis process takes advantage of the meaning expressed by semantics defined in the ontology and the discovery of new knowledge by means of the inferring mechanism of the semantic web. The main advantage was using other learning knowledge inferred on students and domain knowledge to introduce artificial intelligence in decisions making in learning environments. An experimental study using i-SoLearn indicates that this approach is efficient in analysing students' activities and showed significant improvements in the design of the intelligent learning systems.

References

[Aguiar, 2014] Aguiar, E., Ambrose, G. A., Chawla, N. V., Goodrich & V. Brockman, J.: Engagement vs Performance: Using Electronic Portfolios to Predict First Semester Engineering Student Persistence. JLA, 1(3) (2014), 7-33.

[Berners-Lee, 2001] Berners-Lee, T. & Fischetti, M.: Weaving the Web: The original design and ultimate destiny of the World Wide Web by its inventor. DIANE Publishing Company, Darby, PA USA, (2001).

[Brusilovsky, 2003] Brusilovsky, P. & Peylo, C.: Adaptive and intelligent web-based educational systems. International Journal of Artificial Intelligence in Education (IJAIED), 13 (2003), 159-172.

[Carreno, 2014] Carreno, I. D. V. G.: Theory of connectivity as an emergent solution to innovative learning strategies. American Journal of Educational Research, 2(2) (2014), 107-116.

[Chatti, 2010] Chatti, M. A., Agustiawan, M. R., Jarke, M. & Specht, M.: Toward a personal learning environment framework. International Journal of Virtual and Personal Learning Environments (IJVPLE), 1(4), (2010), 66-85.

[D'mello, 2012] D'mello, S. K., Craig, S. D., Witherspoon, A., Mcdaniel, B. & Graesser, A.: Automatic detection of learner's affect from conversational cues. User modeling and user-adapted interaction, 18(1-2), (2008), 45-80.

[Downes, 2012] Downes, S.: Connectivism and connective knowledge: Essays on meaning and learning networks. networks. National Research Council Canada, (2012).

[Fancsali, 2011] Fancsali, S. E.: Variable construction for predictive and causal modeling of online education data. in Proc. 1st Int. Conference on Learning Analytics and Knowledge, Banff, AB, Canada (2011), ACM, 54-63.

[Ferguson, 2015] Ferguson, R. & Clow, D.: Consistent Commitment: Patterns of Engagement across Time in Massive Open Online Courses (MOOCs). JLA, 2(3), (2015), 55-80

[Fortunato, 2010]. Fortunato, S.: Community detection in graphs. Physics reports, 486(3-5), (2010), 75-174.

[Gros, 2016] Gros, B.: The design of smart educational environments. Smart Learning Environments, 3(1), (2016), 15.

[Guité, 2004] Guité F.: Le connectivisme (néo socioconstructivisme) [The connectivism (neo socioconstructivism)], 2004, available at: http://www.francoisguite.com/2004/12/le-connectivisme-neo-socioconstructivisme/ (accessed 12 December 2018).

[Halimi, 2014] Halimi, K., Seridi-Bouchelaghem, H. & Faron-Zucker, C.: An enhanced personal learning environment using social semantic web technologies. Interactive Learning Environments, 22(2), (2014), 165-187.

[Joksimovic, 2015] Joksimović, S., Dowell, N., Skrypnyk, O., Kovanović, V., Gašević, D., Dawson, S. & Graesser, A. C.: How do you connect?: analysis of social capital accumulation in connectivist MOOCs. In Proc. 5th Int. Conference on Learning Analytics and Knowledge, Poughkeepsie, NY, USA (2015), ACM, 64-68.

[Khribi, 2008] Khribi, M. K., Jemni, M. & Nasraoui, O.: Automatic recommendations for e-learning personalization based on web usage mining techniques and information retrieval". In Proc. 8th Int. Conference on Advanced Learning Technologies, Cantabaria, Spain (2008), IEEE, 241-245.

[Kimmerle, 2015] Kimmerle, J., Moskaliuk, J., Oeberst, A. & Cress, U.: Learning and collective knowledge construction with social media: A process-oriented perspective. Educational Psychologist, 50(2) (2015), 120-137.

[Ley, 2016] Ley, T., Klamma, R., Lindstaedt, S. & Wild, F.: Learning analytics for workplace and professional learning. In Proc 6th Int. Conference on Learning Analytics & Knowledge, Edinburgh, United Kingdom (2016), ACM, 484-485.

[Martin, 2013] Martin, T., Aghababyan, A., Pfaffman, J., Olsen, J., Baker, S., Janisiewicz, P. & Smith, C. P.: Nanogenetic learning analytics: Illuminating student learning pathways in an online fraction game. In Proc. 3rd int. conference on learning analytics and knowledge, Leuven, Belgium (2013), ACM, 165-169.

[Martinezmaldonado, 2015] Martinez-Maldonado, R., Pardo, A., Mirriahi, N., Yacef, K., Kay, J. & Clayphan, A.: LATUX: an Iterative Workflow for Designing, Validating and Deploying Learning Analytics Visualisations. JLA, 2(3) (2015). 9-39.

[McAfee, 2012] McAfee, A., Brynjolfsson, E., Davenport, T. H., Patil, D. J. & Barton, D.: Big data: the management revolution. Harvard business review, 90(10) (2012), 60-68.

[Montalvo, 2018] Montalvo, S., Palomo, J. & de la Orden, C.: Building an Educational Platform Using NLP: A Case Study in Teaching Finance. J.UCS, 24(10) (2018), 1403-1423.

[Passin, 2004] Passin, T. B.: Explorer's guide to the semantic web. Manning Publications, Shelter Island, NY, USA, 18 (2004).

[Rego, 2010] Rego, H. Moreira, T., Morales, E. & Garcia. F. J.: Metadata and Knowledge Management drivenWeb-based Learning Information System towards Web/e-Learning 3.0. Int. Journal of Emerging Technologies in Learning (iJET), 5(2) (2010), 36–44.

[Rubens, 2012] Rubens, N., Kaplan, D. & Okamoto, T.: E-Learning 3.0: anyone, anywhere, anytime, and AI. In International Conference on Web-Based Learning, Berlin, Heidelberg (2012), Springer, 171-180.

[Siemens, 2012] Siemens, G. & d Baker, R. S.: Learning analytics and educational data mining: towards communication and collaboration. In Proc. 2nd int. conference on learning analytics and knowledge, Vancouver, BC, Canada (2012). ACM, 252-254.

[Snow, 2015] Snow, E. L., Allen, L. K., Jacovina, M. E., Crossley, S. A, Perret, C. A. & McNamara, D. S.: Keys to Detecting Writing Flexibility Over Time: Entropy and Natural Language Processing. JLA, 2(3) (2015). 40-54.

[Spector, 2014] Spector, J. M.: Conceptualizing the emerging field of smart learning environments. Smart learning environments, 1(1) (2015), 2.

[Vega-Gorgojo, 2015]. Vega-Gorgojo, G., Asensio-Pérez, J. I., Gómez-Sánchez, E., Bote-Lorenzo, M. L., Muñoz-Cristóbal, J. A. & Ruiz-Calleja, A.: A Review of Linked Data Proposals in the Learning Domain. J.UCS, 21(2) (2015), 326-364.

[Verbert, 2013] Verbert, K., Duval, E., Klerkx, J., Govaerts, S. & Santos, J. L.: Learning analytics dashboard applications. American Behavioral Scientist, 57(10) (2013), 1500-1509.

[Wheeler, 2009] Wheeler, S.: E-learning 3.0: Learning through the eXtended Smart Web. In National IT Training Conference. Dublin, Ireland (2011). Retrieved December 12, 2018. http://ittrainingconference.files.wordpress.com/2011/04/ittc_stevewheeler_smartweb.pdf