Recommendations-based on semantic analysis of social networks in learning environments

Abdelaziz Khaled, Samir Ouchani, Chemseddine Chohra

Contents lists available at ScienceDirect

Computers in Human Behavior

journal homepage: www.elsevier.com

Full length article

Recommendations-based on semantic analysis of social networks in learning environments

Abdelaziz Khaled[a], Samir Ouchani[b, *], Chemseddine Chohra[c]

[a] GeantSoft, 75013, Paris, France
[b] LINEACT, Laboratory of Digital Innovation, CESI Research, School of Computer Engineering, CESI eXia, Aix-en-Provence, France
[c] Department of Computer Science, University of Guelma, 24000, Algeria

ARTICLE INFO

Keywords:
Web semantics
Web 2.0
Social network analysis
Social learning networks
User's behavior
Community detection

ABSTRACT

The expansion of social utilization within web applications provides a set of social media that allows users to contribute freely and interact with each other. Consequently, E-learning systems have benefited greatly from the concepts of the social web and the emerging technologies of the web semantic, where students and teachers are living and interacting mainly in the world of Web 2.0 and social networks. Especially, the Web semantic provides content that is understandable by machines and Web 2.0, which help to develop collaborative services. Further it can be integrated in a single social structure semantics that could be a good tool to improve the quality of learning. Our goal is to show that these technologies can be adopted to improve the e-Learning user experience and to provide a full automatic learning platform called iLEARN. We present a learning environment formulated as a social network, in order to carry out an automatic semantic reasoning including the interactions between users as well as their relationships with the provided learning resources. We merge the analysis of social networks with the web semantics to go beyond the analysis of social graphs by integrating a treatment semantics of the knowledge that they contain and designed by the ontology formalism. iLEARNdevelops two ontologies, the first helps to understand the feeling of users versus resources and recommendations, whereas the second categorizes the different resources. In particular, we present an interactive method of detecting communities to provide students belonging to the same learning community the best learning strategies, the strong collaborators candidates, and the relevant resources that meet well within their needs.

1. Introduction

1.1. Motivations

Ever since the beginning of human existence, humans have been organized together by nature at the rate of the need to work and live in a community, to create groups of individuals according to political, social, religious or professional opinion, these groups have evolved over time with the development of them self being and the development of society in all its social and human aspects. The development of these groups gave birth to social networks (Scott, 2000). They can be defined by a social group of people linked by so-called social relationships to achieve a specific goal, purpose, point of view or aspect expressed by the need to communicate with others. Toward the development of human life, social networks are integrated in all eras like: professional (workplace and business), personal life (friendship networks), and learning (schools and colleges).

Social networks are strongly influenced by web technologies and the development of communication tools, they have taken a big part of the lives of people all over the world, social networks have changed their living style as well as their way of communication. The need to exploit this mass of information has created a new need to study the complex structure of social networks, which has led to the emergence of social network analysis science (SNA) (Scott, 2000).

Unfortunately, the web was built for human consumption, not for the consumption of the machine, although everything on the web is readable for the machine, it is not understandable, so a great need to use the web semantic is necessary to express the information in a precise form, interpretable by the machine, ready to software agents to process, share and reuse them, as well as understand the meaning of

* Corresponding author.
Email addresses: Abdelaziz.khaled@geantsoft.com (A. Khaled); souchani@cesi.fr (S. Ouchani); chohra.chemseddine@univ-guelma.dz (C. Chohra)

https://doi.org/10.1016/j.chb.2018.08.051
Received 31 March 2018; Received in revised form 3 July 2018; Accepted 27 August 2018
Available online xxx
0747-5632/

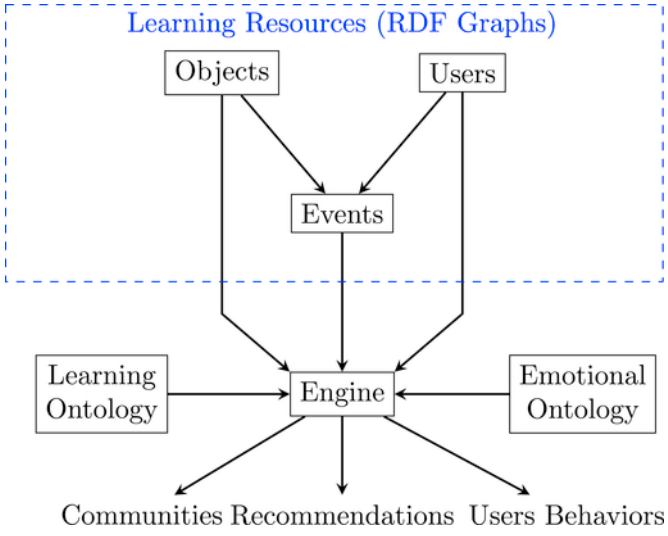

Fig. 1. The iLEARN Framework Schema Overview.

the used vocabulary terms. This would allow mobile/web-based applications to interact both syntactically and semantically (Heflin & Hendler, 2001) between users and resources. Further the concept of Social Semantic Web (SSW) (Gruber, 2008) came into being after the fusion of the best of the social web and the semantic web. It denotes the development in which social interactions on the Web leads to the creation of an explicit representations and semantically rich and sound knowledge.

1.2. Contributions

Our ultimate goal is to develop a scalable and automatic learning environment called iLEARN. It is based on the semantic analysis of the social interactions that takes place between users-users and users-resources, the analysis is based on detecting communities of the semantic data that represents the users and their areas of interest. Further, it provides automatically the different recommendations with respect to the users interests, as well it proposes interactions with other users in the system where iLEARN ensures that these recommendations will certainly improve the learning process by providing students with the best learning practices, the desirable collaborators, and the relevant resources that fit better their needs. To automate the learning process and to help learners and machines to communicate and exchange correctly the semantics of terms, we develop two ontologies (learning and emotion) represented in a meta-data schema that collect a set of vocabulary terms and define their relationships semantically.

iLEARN aims to improve the learning process by taking into account learners' needs, knowledge and preferences, as well to integrate student

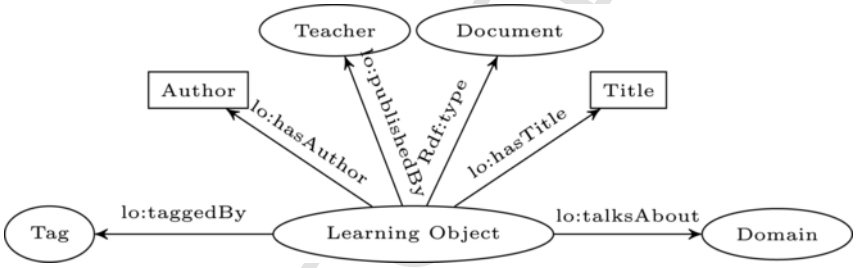

Fig. 2. Learning resource RDF graph.

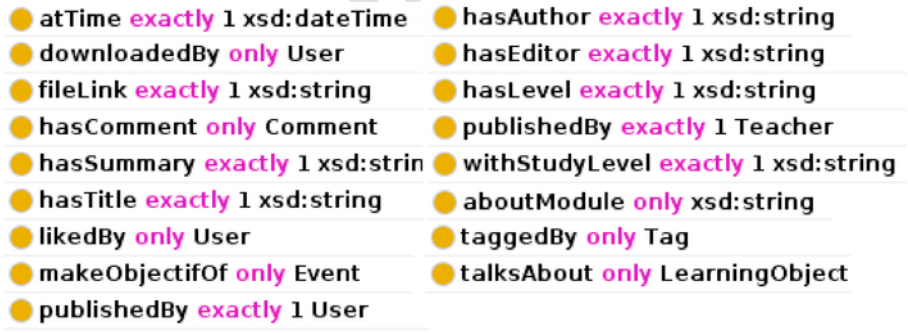

Fig. 3. The learning object properties.

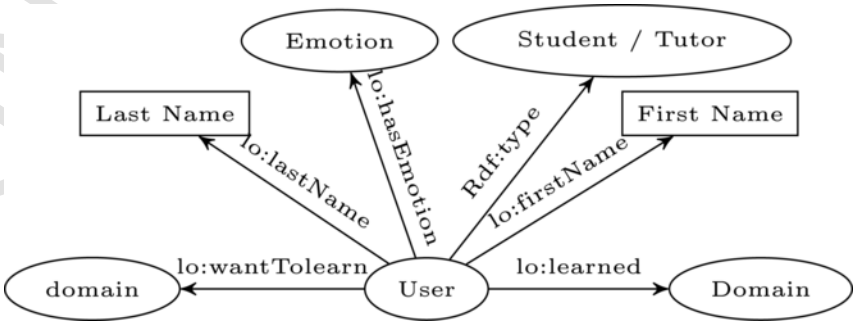

Fig. 4. The user RDF graph.

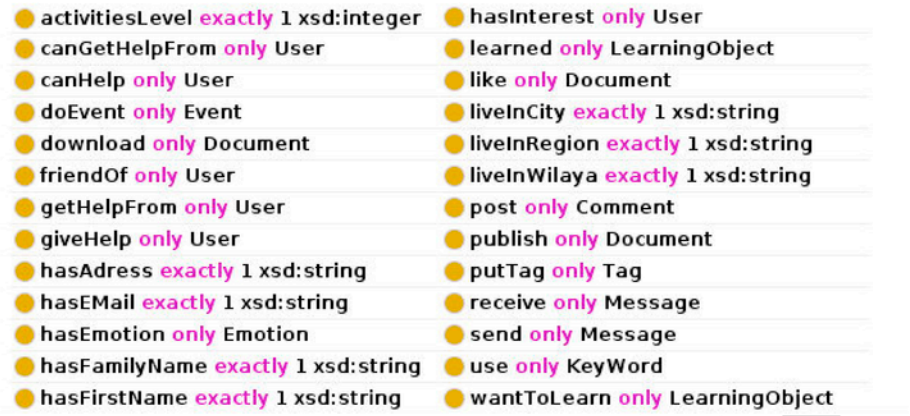

Fig. 5. The user's properties.

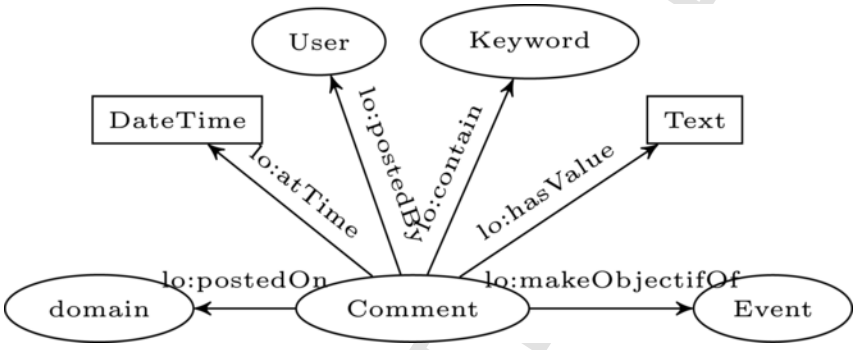

Fig. 6. The Comment RDF graph.

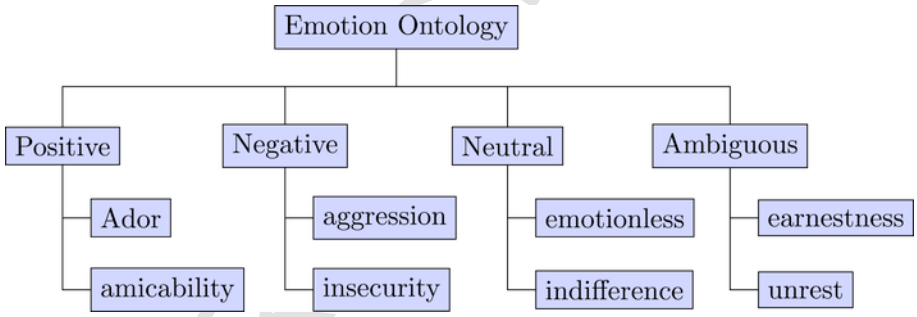

Fig. 7. An extracted part from the ontology of emotions.

Table 1
The most useful terms in the ontology of emotion.

Term	Class
Interested	Positive
Happy	Positive
Bored	Negative
Annoyed	Negative
Tired	Negative
Contented	Positive
Calm	neutral
Nervous	Negative
Good	Positive
Sad	Negative
earnestness	ambiguous
unrest	ambiguous

into the learning community. Our main contributions are described as follows:

- Proposing a fully automatic learning framework, based on web semantics, called iLearn.
- Detecting and grouping the learning communities, and extracting their common interests and the interests of their individual users.
- Providing intelligent recommendations, based on the analysis of the interactions between users and communities and extracting their interest in the network.
- Producing an automated mechanism to capture the understanding semantics of the users feelings and learning needs based on two ontologies: emotional and learning resources.
- Guiding learners by providing them the best learning paths in accordance with the recommendation of the best learning resources.

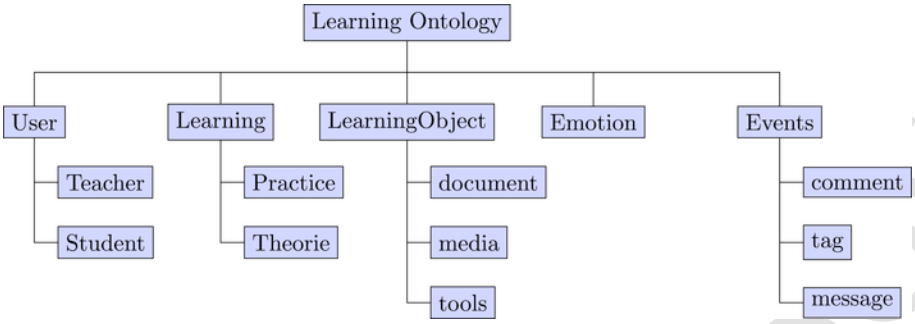

Fig. 8. An extracted part from the learning ontology.

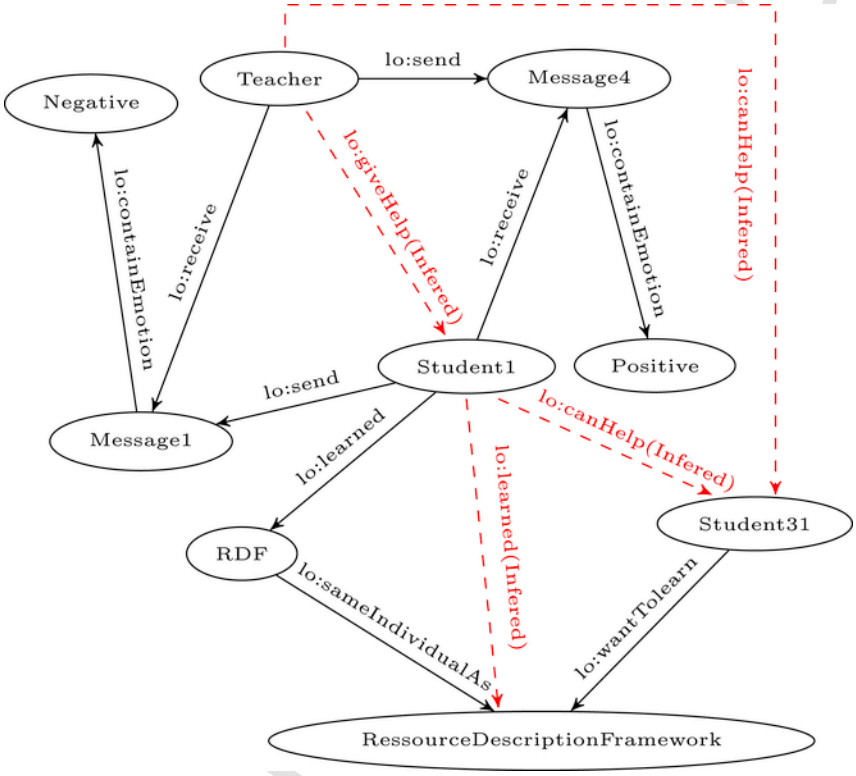

Fig. 9. Inference mechanisms in web semantic.

1.3. Outlines

In the following, we review the related work in Section 2. Then, Section 3 develops the theory and techniques behind iLEARN to model and analyze a learning system. In section 4, we describe the developed tool and Section 5 shows and discusses the obtained experimental results. Finally, Section 6 concludes the paper and sketches directions for future research.

2. Related work

In this section we review the existing initiatives that treat the analysis of social networks and models the users comportment using ontologies in order to show how well our contribution advances the existing techniques in crowd and anticipatory computing especially in the social networks and learning eras.

The Analysis of Social Networks (ASN) is one of the main prominent approaches that classify and analyze the user's behavior (Mansur & Yusof, 2013) at many stages and levels of abstraction especially they are relying on graph specification where nodes and edges represent the users and their relations respectively. In literature, many initiatives are dedicated to study ASN quantitatively and qualitatively. According to Scott (2000), density is a measure that can be used in the context of a global network or an individual centered analysis, which consists to calculate the density of links related to a given node in a graph. Another concept is the centroid of a social network that defines what can make a node more central and important than another (Freeman, 1979). In this direction, several approaches are proposed such as centrality of degree (Nieminen, 1974) that considers central nodes having the highest degrees in the graph. Further, centrality of intermediation (Freeman, 1979) that focuses on the ability of a node to serve as an intermediate one in a graph. Furthermore, the centrality of proximity (Freeman, 1979) is to evaluate the centroid of a node by looking to the size of paths that bind a given path to another ones.

Another important concept, especially in crowd intelligence, is the detection of communities that has been recognized as one of the main important issues in ASN. Qi et al. (2014) proposed an approach to group communities by assigning densities to each relation representing the parent-child pairs. For the density evaluation, it applies the QCM algorithm that takes a graph as input and produces a Dendrogram as output. The important feature of this method is the overlapping be-

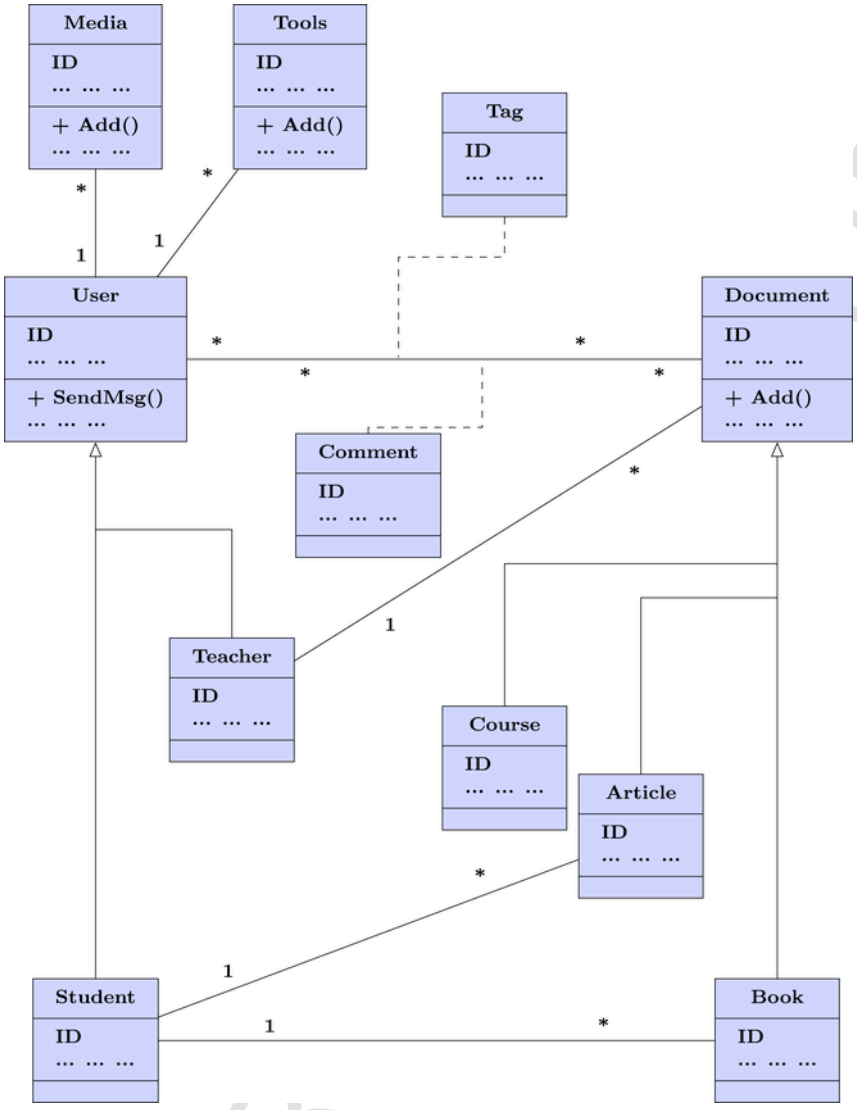

Fig. 10. The class diagram for iLEARN system.

tween clusters with a multi-membership values (Ou and Zhang). Compared to our approach, Qi et al. (2014) exclude the semantics analysis of the users relationships as well as their behaviors and intentions. Especially, they construct a weighted graph from a Dendrogram then applying max-flow and min-cut on the new obtained weighted graph. Further, they evaluate their approach with a couple of experiments on computer generated graphs where the community structure is already known and it does not need an automatic detection of communities.

Sun & Sun (2017) developed a framework to detect communities on a complete graph that models a separated groups (communities). They divided the network into subnetworks by maximizing the difference between the interior and the exterior of subnetworks when the exterior is a complete bipartite graph. To experiment their approach, they validate the framework on computer-generated networks. However, the authors have not discussed what is the meaning of the relations between nodes or what is the common points between nodes of a given community which is the main objective in the current work and crowd intelligence.

Martinez, Dimitriadis, Rubia, Gómez, & de la Fuente (2003) proposed a qualitative method for ASN to study collaboration and participation of students in the classroom. They presented an evaluation in a university course of Computer Architecture during two academic years. Their study was done with a text analysis tool that allows an automatic processing of computer logs by social network analysis, as well as the qualitative research tool Nud*IST.[1] The experimented data set comes from tow sources, the first is from a computer based tool that students use to fulfill the course requirement (event logs), while the second part is collected by traditional means (formal observations, questionnaires). The first step uses EL2AM (Event Logs to Adjacency Matrices) to transform the event logs into adjacency matrices that represent a social network, then they calculate the density and degree of centralization. As second step, they used NUD*IST for qualitative processing. The main weaknesses of this contribution is not proposing recommendations either their impacts after application.

Dráždilová, Martinovič, Slaninová, & Snášel (2008) have used the agglomerative hierarchical clustering method to extract the social network communities from Moodle E-learning system by computing their similarity matrix. The student in Moodle is represented by a vertex in a social network and the relationship is defined by his activities in Moodle. When two students have the same activities, automatically they have a strong relationship between them. Initially, they define the activity sequences for each student in relation to each subject *'resource*

[1] NUD*IST is a robust and reliable program, appropriate exclusively for qualitative research analysis.

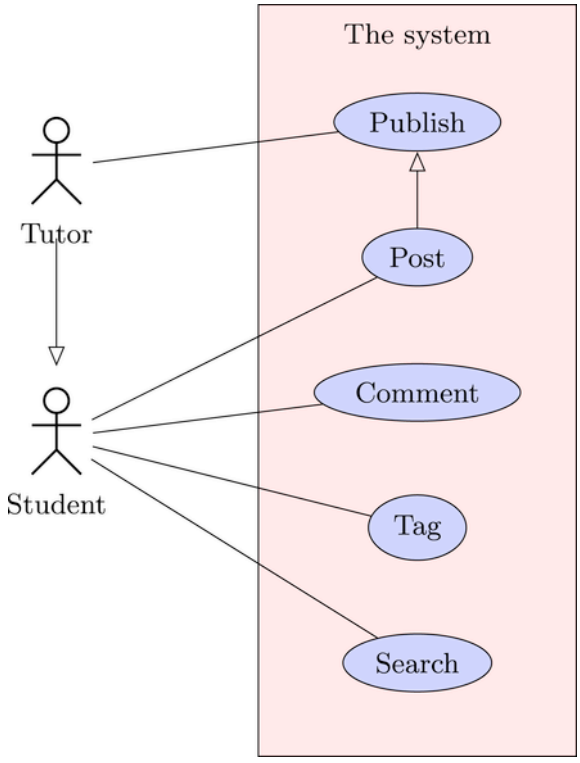

Fig. 11. The use case diagram for iLEARN system.

view', *'forum view'*, *'choice view'*, etc. Then, they generate a matrix that represents the activity types for each student and define the similarity by calculating the Cosine measure in order to apply the agglomerative hierarchical clustering method that extracts the communities. In general, they focused only on the number of activities to represent the relationship between students which limits the generation of communities of a student has different center of interests.

Aşkar (2011) proposed ASN encapsulating an analysis chain-tool to investigate the relations between learners, teachers, and scientists by taking into account the naturalistic or self-reported data. Askar's ASN looks only on the distance between nodes without considering the meaning of the relationship across users.

The discussed literature shows that ASN has been studied deeply by exploring different techniques. However there are some works consider ontologies as a representation formalism in e-learning before it has been analyzed by ASN. Iván et al. (Cantador & Castells, 2006) proposed an approach that clusters the ontologies in a multi-layered social network by comparing the user profiles of interest in order to find the similarity between users. They proposed a three-steps approach, spreading the semantics of the user preferences, extending the initial sets of individual interests, and clustering the semantic domain concepts into concept groups, based on the vector space of user preferences. Finally, users are clustered in order to identify the closest class to each user.

Yusof, Mansur, & Othman (2011) and Li, Yang, He, & Ai (2010) proposed a new paradigm to analyze social networks using the ontology concept. Yusof et al. (2011) used the ontology to give a semantic meaning of the relationships among users. This ontology is designed to capture the learning patterns of students, teachers and other elements by using two kinds of relationships: *Knowledgeable* expresses the communication between users (discussion about the course/subject) and *Communicable* that models the social interactions without any restriction to a specific course or subject topic. Li et al. (2010) divided the relationship into two relations: Enemy and Kind. The latter has been divided into sub-relations like friend, cooperate, etc. However, both contributions have not discussed how to analyze the network with the ontology or how it can improve the analysis of social networks. Further both ontologies cover a precise sample of terms and relations.

Halimi & Seridi-Bouchelaghem (2015) proposed a social personal learning environment based on the results presented by Khaled et al. (Chohra & Khaled, 2013) to improve a personalized learning process using the social web semantic and the learning styles where the analysis has been done through Bayesian networks. Their proposed approach consists of three steps: utilizing web semantic modeling techniques (Ontology, RDF graph) to design knowledge and to provide learners a precise content, MBTI test has been applied to determine learner's style and to generate their appropriate content, as well Bayesian network to filter and validate the resources content depending the learner's learning styles. However, the authors have not discussed how to analyze the user traces and how to generate a user behavior which is the main evident point in social learning systems that we target and focus on in the present work. Further they do not generate the recommendation as proposed by iLEARN.

Obeid, Lahoud, El Khoury, & Champin (2018) discuss an approach for developing ontology-based recommender system based on machine learning techniques to orient students in higher education. The proposed recommender system is an assessment tool for students' vocational strengths and weaknesses, interests and capabilities. The proposed ontology-based recommender system is to identify the student requirements, interests, preferences and capabilities to recommend the appropriate major and university for each one. The status of this work is in its first stage where the ontologies architecture are discussed only and not developed, also the proposed analysis based on machine learning is not tested yet.

Recently, Jetinai (2018) proposed a resource recommendation approach based on reasoning rules to personalize the on-line learning for a specific domain/category. K. Jetinai designed a reference ontology that describes the learning style to each learner. Also, he proposes a personalization rules to support personalized semantic search for heterogeneous learning resources. In fact, personalizing such inference rules and ontologies are not clarified, and/or they are not specified for a precise group. Compared to our framework, iLEARN is applicable for any domain, extensible and scalable for any changes in terms of rules inference or ontology development.

Based on the reviewed initiatives we believe that our proposed contribution is scalable, automatic and generic that can be personalized to any learning domain. Further, we think that this paper helps to improve the state-of-the-art in crowd and anticipatory challenges, and it can be considered as fundamental for further future research directions.

3. iLearn framework

iLEARN framework analyzes social networks in a learning environment through three main components as shown in Fig. 1, the learning resources, the ontologies, and the analysis engine. The learning resources are composed from objects, users, and their events that are described in Resource Description Framework (RDF), a standard language for describing resource meta-data, as RDF graphs. The ontologies represent the learning domains and the user's emotions. The engine takes into consideration the ontologies and the RDF graphs of the resources to generate the behavior of users, their communities, and analyze them to provide the appropriate recommendations.

3.1. The learning resources RDF graphs

The RDF representation of a given learning resource is a rooted directed graph expressed by the tuple ⟨*subject*,*predicate*,*object*⟩ where the edges link the root the nodes. The root *subject* is the resource that we want to describe and it is identified by URI (URI *'Unifrom Resource Identifier*' is a string characterizing a resource). The edges *predicate*, also called properties, are the different types of properties related to the

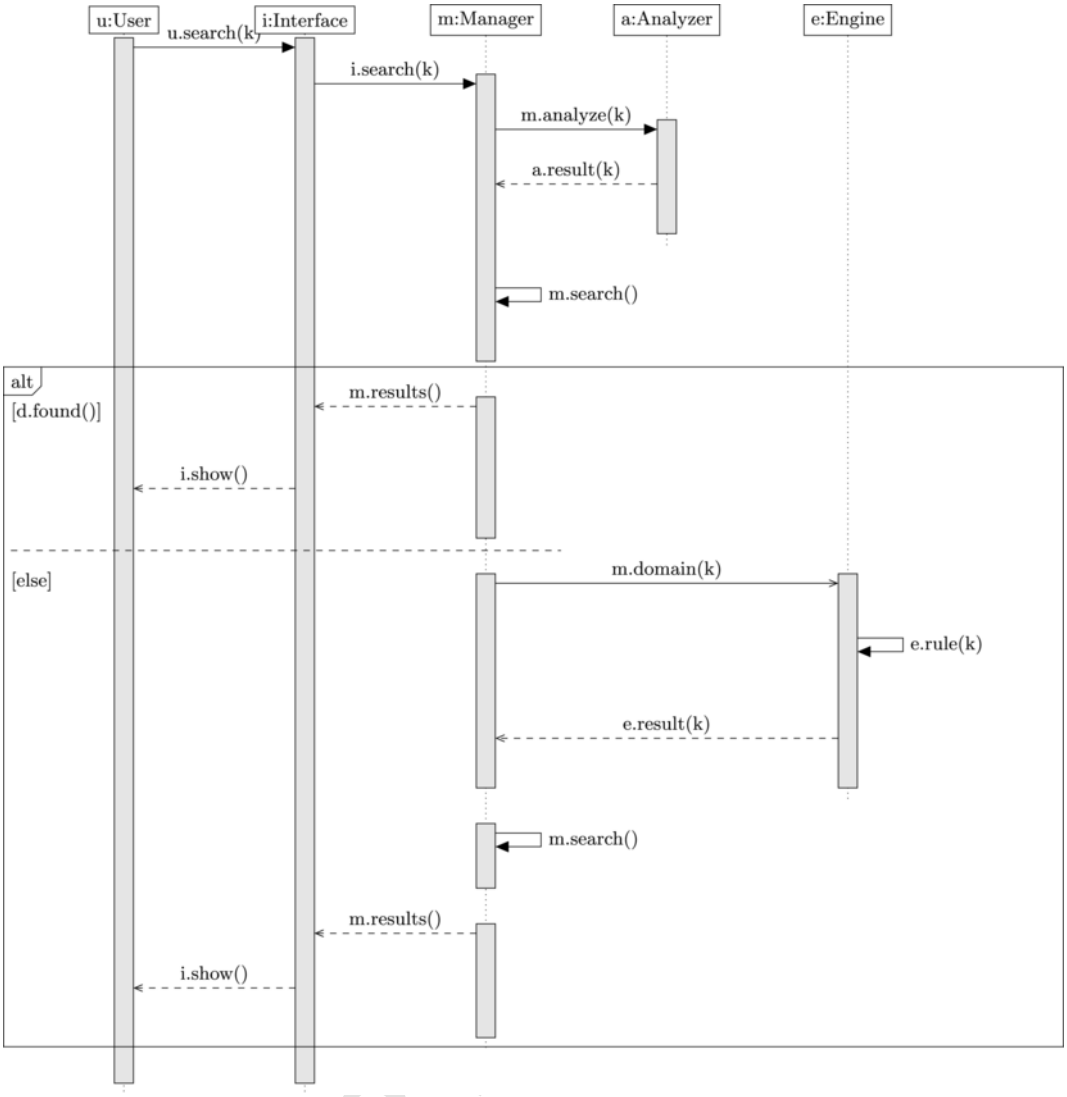

Fig. 12. The sequence diagram for searching documents.

subject, and they are identified by URI. The nodes *object* are the values of the predicate that can be another resource or a literal value. For example, a book is described by: $\langle lo : book1,\ lo : hasTitle, semanticweb\rangle$, where *lo* is an URI, *subject* is a book, the predicate *hasTitle* is a property in the ontology and the *semantic web* is a literal value.

To describe and represent the learning domain on the web semantic, the used learning words (terms) define the main concepts and relationships in the learning domain. They also help data integration when ambiguities may exist on the used terms that also may lead to integrate new inappropriate relationship. To avoid these issues, we have to describe and define different forms of terms in a standard format such as RDF. Especially the latter offers taxonomic relations, object relations and data-type properties in a standard format. By using RDF, the analysis become more meaningful since it considers the relationship between terms and objects, also it will be aligned within the existing standard of web semantics and extensible for other application domains. In the following, we propose the RDF graphs related to the learning resources: the learning objects, users, and events.

3.1.1. Learning object

A learning object is used by the users to learn a specific topic and it can be a document, video, audio, a tool, etc. For a better visualization, Fig. 2 captures the interesting parts of the RDF graph dedicated to a learning object (It is extracted from a larger refined graph).

Further, Fig. 3 illustrates the main properties describing a learning object. For example the properties, *lo:talksAbout* and *lo:taggedBy*, help the system to understand well the content of the learning object. The keyword *only* denotes *Universal restrictions* and *exactly* is to create an exact cardinality restriction.

3.1.2. Users

Initially, iLEARN supports two main categories of users, students and tutors, where both of them can execute similar actions except tutors that can publish courses. The tutor can play the role of a student but not the inverse. The RDF graph presented in Fig. 4 shows a selected features of objects and abilities in terms of predicates of a user that is considered as subject in this case of RDF graph. A user is identified by his name, emotion, and position, as well he has a background in a given domain but he wants to learn other domains and/or topics.

The user model contains the knowledge that the system needs about him/her. iLEARN models the user with different properties, *lo:learned* presents the knowledge of the user on a specific learning domain, *lo:wantToLearn* expresses the needs of the user in terms of learning, *lo:giveHelp* and *lo:getHelpFrom* are irreflexive properties between users

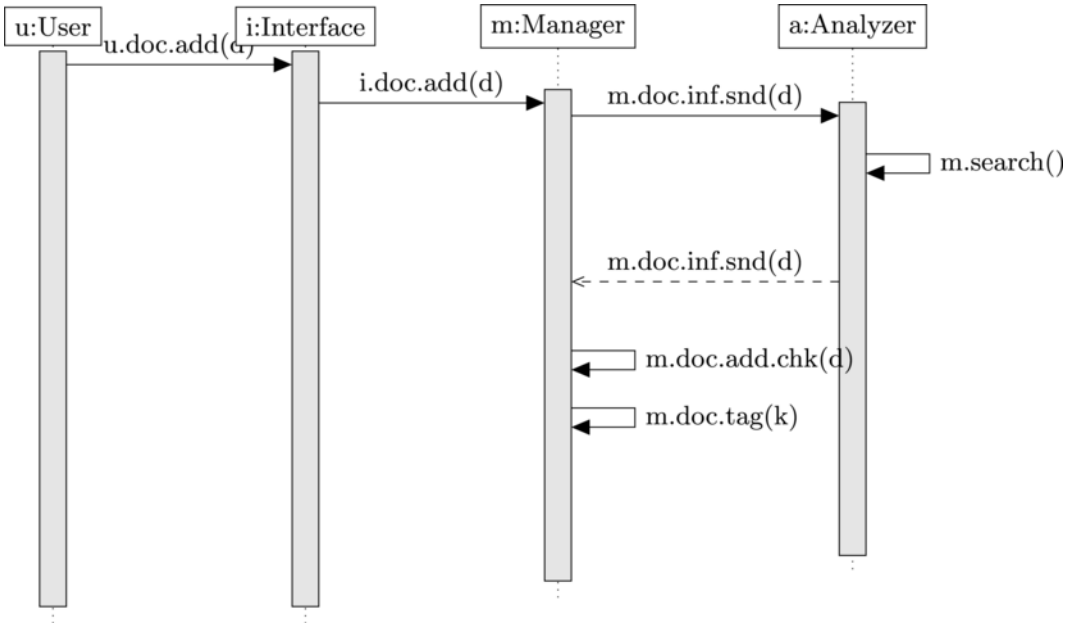

Fig. 13. The sequence diagram to publish a document.

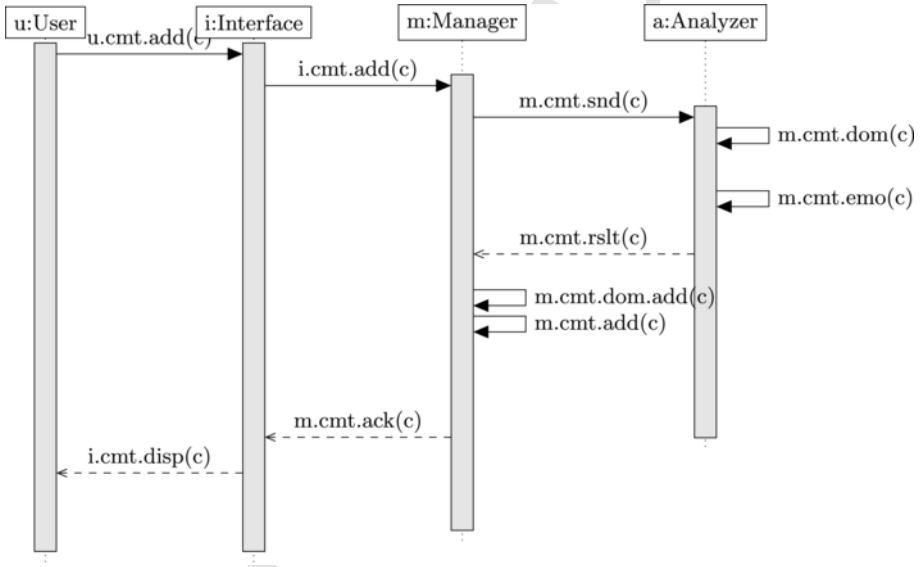

Fig. 14. The sequence diagram to add a comment.

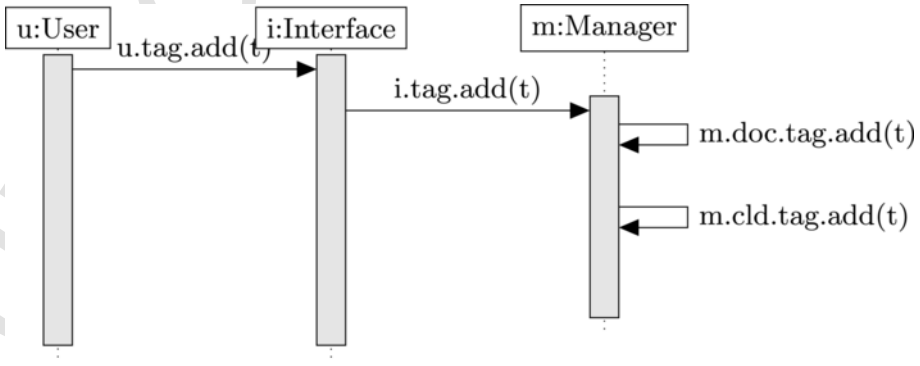

Fig. 15. The sequence diagram to add a tag.

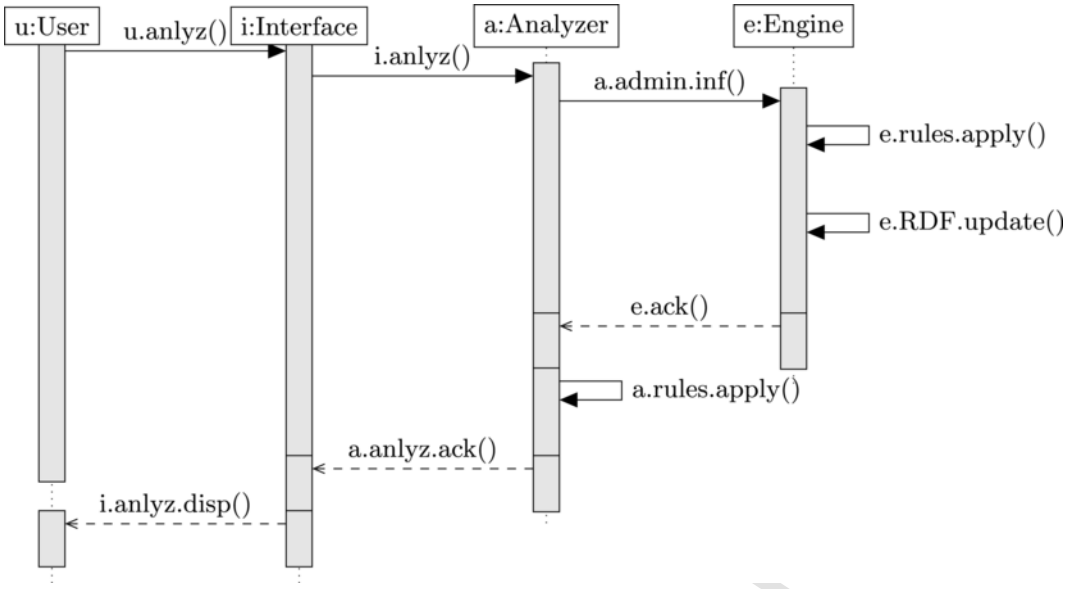

Fig. 16. The sequence diagram for the administration analysis.

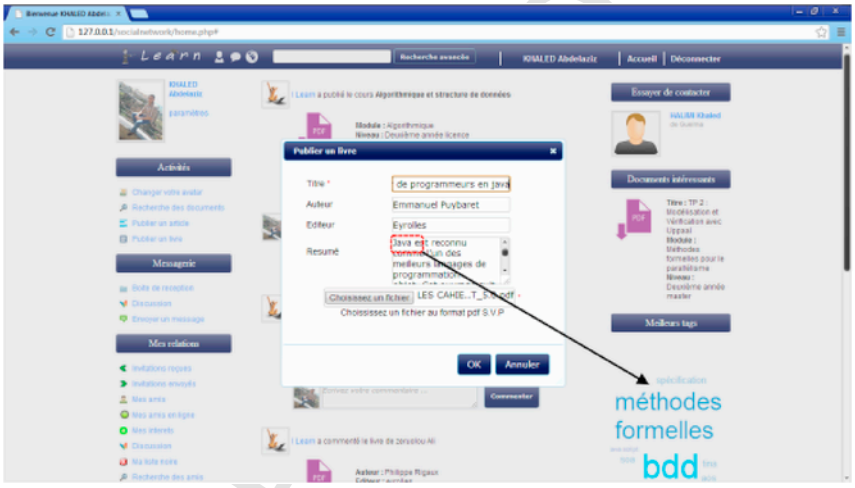

Fig. 17. Main interface of iLEARN system.

to express if a user can get help from a user or give it to another one. Fig. 5 lists the different properties assigned to a user.

3.1.3. Events

The events explore the users' behaviors and their feelings regarding a learning domain. More precisely, the events of a user is the history of actions executed by him/her. Fig. 6 shows the RDF graph of a comment as a single action by a given user.

3.2. The ontologies

For a full automation of the framework, we rely on ontologies that allow to describe formally the different knowledge of users, resources, emotions, tags and the communication tools. Ontolgies in iLEARN allow us to:

- Give a concrete meaning and describe the learning resources.
- Define formally the learning domain.
- Define formally the emotions of the users.
- Improve the recommendation system.

The ontology describes two types of knowledge: concept and relationships, the former represents a variety of things like person, books, etc whereas the latter is a way in which two concepts can be connected to each other. Basically, the ontology sounds like the concept of RDF, but the difference is that the ontology has a collection of logical rules.

Basically, iLEARN develops two ontologies, the ontology of emotions that is used to understand the feelings and needs of users whereas the learning ontology covers the learning domains and their dependencies.

3.2.1. The ontology of emotions

The ontology of emotions presents the different terms and words that can be expressed by the user, as well, it associates for each word a class of emotion. Four classes of emotions are distinguished (*positive*, *negative*, *ambiguous*, and *neutral*). *Positive* emotion means that the user has no difficulty to understand a learning domain and he likes this domain, *negative* allows the system to understand that the user has difficulties with a learning domain, and in this case the system will recommend other learning resources as well as other users who have positive feelings in this domain, *ambiguous* represents a mysterious behavior of the user, and in this case the system keeps analyzing his events to un-

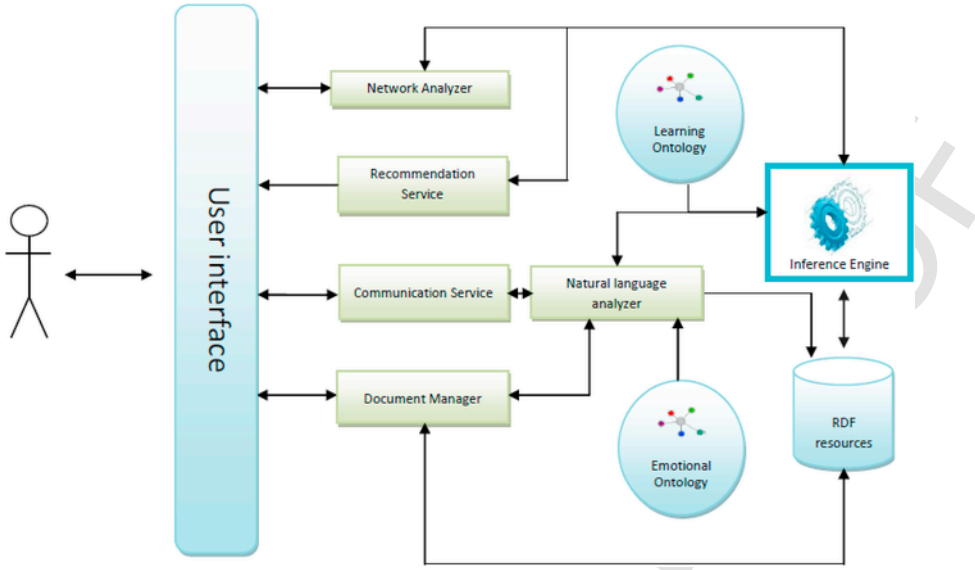

Fig. 18. iLEARN architecture.

Table 2
User's traces on iLEARN.

Action	Number of times	#users/#action	Emotion
Publishing	151	10[4]	
		18[3]	
		17[2]	
		23[1]	
		2 [0]	
Comment	146	20[3]	%66 positive - %34 negative
		30[2]	%47 positive - %53 negative
		13[2]	%45 positive - %55 negative
		7[0]	%0 positive - %0 negative
Tag	163	33[3]	
		21[2]	
		11[2]	
		5[0]	
Message	258	63[> 1]	%57 positive - %43 negative
		7[0]	%0 positive - %0 negative
Search	500	70[> 1]	

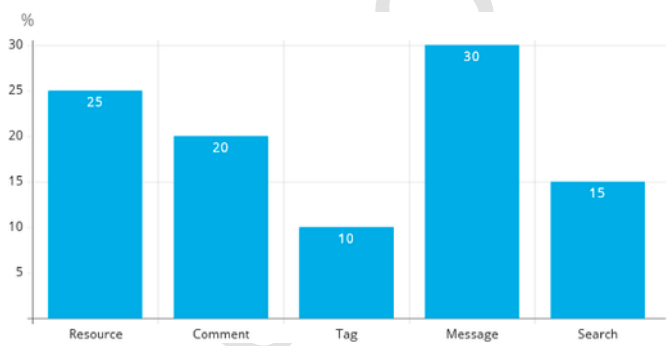

Fig. 19. Overall user's actions.

derstand him better, and *neutral* expresses that the user is not interested to the learning domain. Fig. 7 captures the abstract overview of this ontology.

iLEARN parses each user's posts, comments and messages to find specific words that are defined in the ontology of emotion and associated to at least one of its class. Therefore, iLEARN is capable of understanding the feeling expressed by any user in association with any

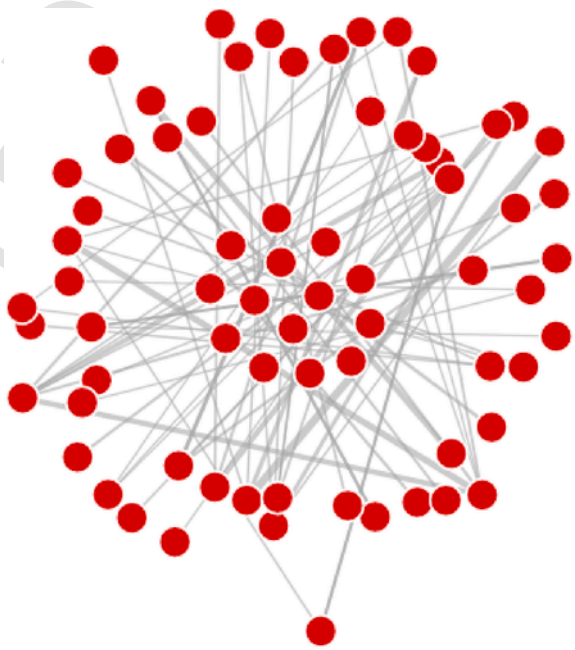

Fig. 20. Social network graph.

event that the latter generates. Table 1 summarizes the most useful words.

The goal behind using this ontology is to understand the user's feelings and to adapt the behaviors of the system according to his emotions. To connect both ontologies, emotional and learning, we added an emotion class in the learning ontology that includes the four main classes from the emotion ontology as an instance of this class.

3.2.2. *Learning ontology*

Compared to the previous ontology, where individuals belonging to a hierarchy of classes, the learning ontology defines different type of users, their related actions in the network including the existing learning domains. Fig. 8 shows the important part of the learning ontology: users as teacher and student, learning if it is practice or theory, type of the learning object, emotion, and the event that can either comments, tags, or messages.

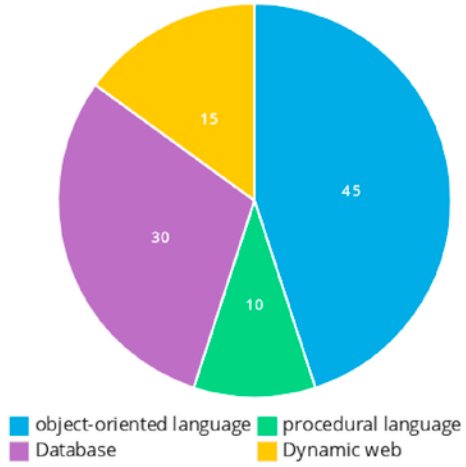

Fig. 21. Detecting the students behaviors.

Fig. 22. Visual representation of users belonging to different communities in ILEARN.

3.3. Engine

The engine consists of two main components: the *Inference* and the *Analyzer*. The former allows the reasoning on the acquired knowledge in the RDF graphs with the rules (ontologies) to deduce the conclusions as results. The latter extracts and analyzes the obtained results to produce the users behaviors, communities, and the appropriate recommendations.

3.3.1. Inference engine

The inference is built-up RAP (Oldakowski, Bizer, & Westphal, 2004) using forward chaining and backward chaining inference algorithms. The forward adds a new inference rule to any new statement from RDFs or a name space, whereas, no inference is done when a new statement is added to the model in the second case. When a query is executed against the model, only the necessary inferences for this query are selected to be checked against an inference-rules index that produces conforming statements.

3.3.2. Analyzer

To generate the user behavior, and/or to group the communities, and/or to produce the recommendations; the analyzer develops for each objective an appropriate solution. Algorithm 1 takes the set of events (comments, tags, likes) as input and generates the normalized matrix V that represents the report of interest for each user with respect to the learning domains i.e; $\sum_{u\in\mathbb{U}}\sum_{d\in\mathbb{D}} v_{u,d} = 1$ where $\mathbb{D}$ is the set of learning domains and $\mathbb{U}$ is the set of users.

Algorithm 1
ExBeh: Extracting the user's behavior.

Require: $\cup$ {set of users.}
Ensure: V {Matrix of the report of interest.}
1: **for** $u \in \mathbb{U}$ **do**
2: $kWords = Match(e.events, Ont)$; {Match the event with ontology.}
3: $v_{u,*} = Inf(kWords, Ont)$; {Infer the report of event for an ontology per domain.}
4: **end for**
6: $V = Normalize(V)$; {Normalize V.}

The function *Match* retrieves the set of keywords from the string *event* by matching the similar words between *event* and the ontology *Ont*. Then, the function *Inf* relies to the ontologies to infer the set of domains related to the keywords based on the semantics of each word in the developed ontology. Finally, the normalized matrix V contains the report of interest for each user to the learning domains; i.e., the number of keywords for each domain by a user profile.

Communities. To extract the communities, Algorithm 2 calculates the similarity between users taking into account the matrix V as input. The similarity between user's is evaluated by the Euclidean distance over the interest ratio values using $Sim : \mathbb{U}\times\mathbb{U} \rightarrow \mathbb{R}$ and $Sim\left(u_i, u_j\right) = \sqrt{\sum_{d\in\mathbb{D}}\left(v\left(u_i, d\right) - v\left(u_j, d\right)\right)^2}$. Then, it generates the communities C by applying K-means clustering algorithm.

Algorithm 2
Detecting the communities.

Require: V,$\cup$
Ensure: C
for $u \in \mathbb{U}$ **do**
for $u' \in \mathbb{U}$ **do**
$S = Sim(V(u,*), V(u',*))$; {Measure the similarity between users.}
end for
end for
$C = Kmeans(S)$; {Retrieve the communities.}

Algorithm 3
K-means algorithm.

Require: $\cup$, k;
Ensure: $C_1,\ldots C_k$; {Centers of clusters}
{Randomly, assign centers to the clusters.}
1: **for** $(All\ C_i)$ **do**
2: c_i *the center of* C_i $=$ *Random*;
3: **end for**
{Assign users to the closest cluster.}
4: **for** $u \in \mathbb{U}$ **do**
5: *Measure the distance between u and each center*;
6: *Assign to u it to the closest cluster*;
7: **end for**
8: **if** (*Changing clusters*) **then**
9: *Recalculate the centers of gravity*;
10: *Go to line* 4;
11: **end if**

Recommendations. The inference mechanisms of the web semantic for creating the recommendations is based on the social communities and the user's needs. ILEARN recommends the learning resources (users, objects, etc) for whom are in the same community. Further, we defined several properties in the learning ontology to recommend the precise resources to the right users depending their needs. The following is a selected set of properties that the ILEARN takes into consideration.

- *wantToLearn(u,d)* means the user u wants to learn the domain d.

Table 3
Statistics on communities generated by Algorithm 2.

Community	C01	C02	C03	C04	C05	C06	C07	C08	C09	C10	C11	C12	C13	C14	C15	C16	C17	C18
#users	7	10	6	3	3	5	2	3	7	1	1	4	4	3	3	3	2	1
mean distance to centroid	52.88	60.40	51.31	49.52	52.48	73.47	65.01	42.43	71.12	0	0	74.86	66.79	70.69	55.97	66.12	53.95	0

Table 4
Statistics on communities generated by Algorithm 2.

	C01	C02	C03	C04	C05	C06	C07
C01	0	265.43	238.82	359.50	285.57	378.76	250.40
C02	265.43	0	255.59	153.66	217.14	253.41	201.32
C03	238.82	255.59	0	301.84	187.77	290.67	178.68
C04	359.50	153.66	301.84	0	187.85	210.43	225.67
C05	285.57	217.14	187.77	187.85	0	245.86	217.74
C06	378.76	253.41	290.67	210.43	245.86	0	267.48
C07	250.40	201.32	178.68	225.67	217.74	267.48	0

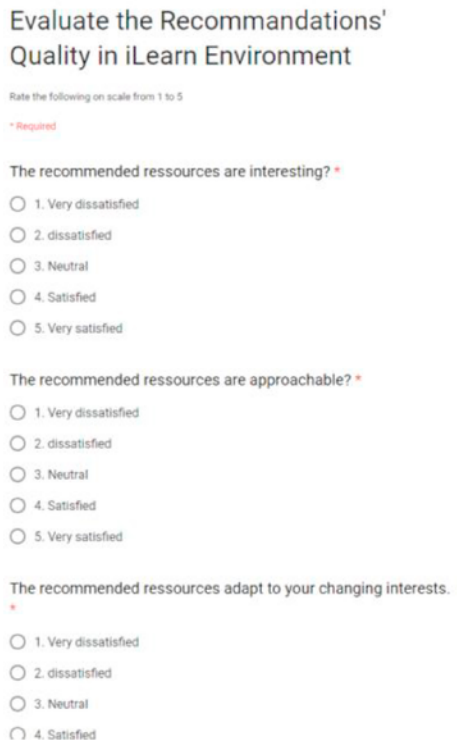

Fig. 23. Questionnaire filled by ILEARN users.

Fig. 24. Recommendation quality based on the users rating.

- *learned(u,d)*; the user *u* masters the domain *d.*
- *CanHelp(u,u',d)⇔wantToLearn(u',d)∧learned(u,d)* is a transitive property that means the user *u* can help the user u' in domain *d.*

For example, Fig. 9 shows how we apply the inference in the recommendation. The relation *lo: giveHelp* between *Teacher* and *Student1* is inferred from the contents of messages (like Message1 and Message4) exchanged between both of them. The relation *lo: learned* was inferred because the student *Student1* masters the learning object "RDF" and the relation between *RDF* and *ResourcesDescriptionFramework* is *owl: sameAs*, i.e; both individuals are similar.

The relation *lo: canHelp* between *Student1* and *Student2* is inferred because the first masters a domain which is a learning objective for the latter. Finally, the relation *lo: canHelp* between the teacher *teacher25* and the student *student31* is inferred because of the transitivity of the relation *lo: canHelp* and also it is a sub property of *lo: giveHelp*.

4. iLearn tool

This section describes the structural and behavioral diagrams of ILEARN. Fig. 10 shows the main classes representing the structure of users and the learning resources in ILEARN. Further, we realize the proposed ILEARN framework by identifying the actors and roles with their needs and functionality.

Schematically, we illustrate them through the use case diagram presented in Fig. 11. The tutor instantiates the role of the student that can publish a document, post, comment, tag and search for a document or a comment using a key word; whereas the administrator instantiates the tutor role with the ability to manage and analyze ILEARN system. The detailed behavior of each actor's operations are described using sequence diagrams as follows.

4.1. Searching operation

The sequence diagram of the searching operation is showed in Fig. 12, and sketched in the following three steps.

- First, the actor indicates the keywords to be used for the searching operation through the user interface.
- Then, the document manager requests the natural language parser to extract the meaning of the keywords based on the ontology presented earlier.
- The document manager applies the searching algorithm on the results returned by the natural language parser. If no results found, it extracts the nearest domains from the inference engine by applying the inference rules, then, the document manager applies again the searching algorithm.

4.2. Publish a document

The sequence diagram showing the operation of publishing a document is depicted in Fig. 13. The execution steps of this operation is described as follows.

- The user enters the document information and upload the file.
- The natural language analyzer processes the introduced document information and extracts its associated learning domains.
- The document manager adds the document and updates the cloud tags.

4.3. Add a comment

To add a comment, Fig. 14 illustrates the execution steps related to this operation that are described here.

- The user writes the comment and the manager sent it to the analyzer.
- The analyzer analyzes the comment by extracting the emotional and the learning domains descriptions of this user regarding the comment.
- The user's learning profile is updated according to the obtained results.

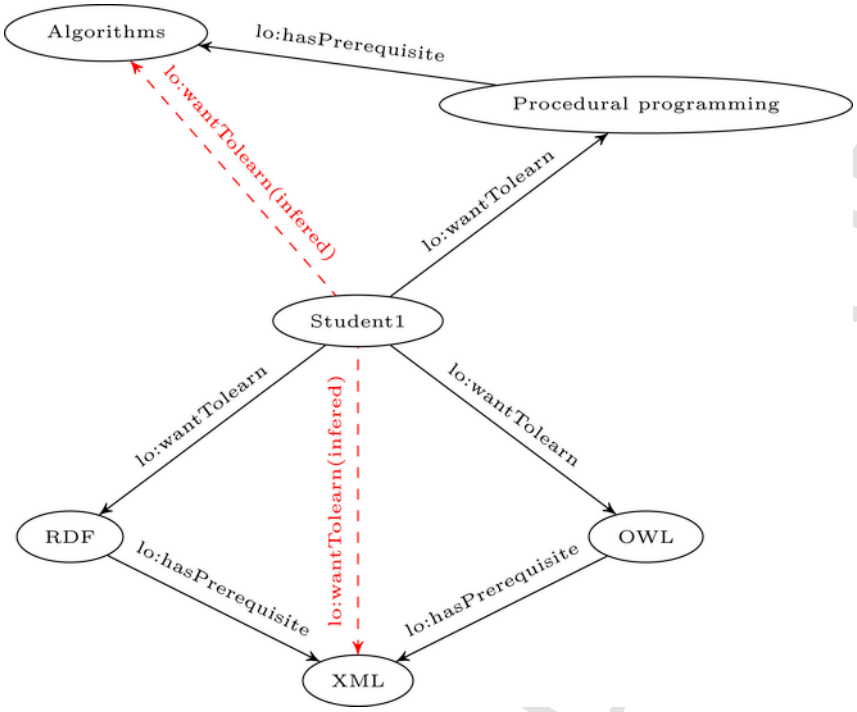

Fig. 25. Inference guided recommendation.

4.4. Adding tags

Fig. 15 presents the steps to add a tag to a comment as a sequence diagram and its related operations are defined as follows.

- Always through the user interface, the user fills the keywords to be added as a tag to an object document/comment.
- Then, the document manager adds the tag and updates the cloud tags.

4.5. Network analysis

The network analysis is a task dedicated to the administration of the system and Fig. 16 displays the sequence diagram to perform a global network analysis by executing the following steps.

- The administrator selects an analysis operation.
- The network analyzer asks the developed engine to apply the rules of inference to make available the new relations which is the main interest of the use of semantic analysis.

5. Implementation and experiment

In this section we experimented our ILEARN system with the student of eXia CESI[2] computer science engineering school. Fig. 17 presents a screen shot of ILEARN system focusing on its various features where Fig. 18 shows the global overview of its architecture.

The experiment was done with 70 students of different levels. Table 2 summarizes the obtained traces of the experiment. It shows the type and the times of the executed actions, the number of users executing the same action, and the ratio of the extracted emotion.

Fig. 19 represents the average of the executed actions for the most active users. The statistics show that exchanging messages is 30%, resources 25%, comments 30%, search 15%, and tags 10%.

5.1. Analysis of the network

Fig. 20 depicts the overall social learning network of CESI eXia using a graph where nodes are users (circles in red) and their relationships are the edges (links in gray).

5.1.1. Detecting the students behaviors

Fig. 21 shows the results of Algorithm 1 which is based on the semantics of the user's traces and allows us to detect communities or to customize the learning process. Object-oriented programming takes advantage of interest for students by 45%, data base is in the second range of interest by 30%. Then, dynamic web by 15% and procedural languages by 10%.

5.1.2. Communities detection

Fig. 22 shows a visual representation of the results generated by Algorithm 2 which also rely on the outputs of Algorithm 1. Users belonging to the same community are presented with the same color, while different colors present different communities.

The principle behind the clustering algorithm that we use for our communities detection is to consider each user as an individual (or an observation), and his interests as corresponding variables. Euclidean distance is used to measure the similarities of two users' interests. Users with similar interests are gathered in the same community.

Table 3 shows the detailed statistics about communities generated by the previous clustering algorithm. The first row shows community id, the second indicates the number of users in each community, while the third row exhibits the mean distance between individuals (users) belonging to this cluster (community) and its centroid. The smaller this number is the more interests in common users belonging to this community have. Note also that there are some communities with a single user, however, for larger networks each user has more chance to find members with similar interests. Therefore, we believe that communi-

[2] https://exia.cesi.fr/.

ties with a single user occur less in larger networks with more interactions.

We show finally in Table 4 the distance between the centroids of communities C01 to C07. Note that the shown distances are almost an order of magnitude larger than those presented in Table 3. Such large values implies that the users close to one centroid have a completely different set of interests than those who are close to another (or belonging to a different community).

5.1.3. Recommendations

In order to provide to the users interesting and divers learning materials, we rely mainly on two recommendation strategies.

The first strategy focuses on the user as a single entity. We check via the relations that are defined in our learning ontology what the user wants to learn, his interests and the prerequisites that he needs to achieve his goals. Finally we look for learning resources that respond to the previous criteria. The second strategy focuses on providing the same learning resources to users in the same community. If one express an interest in some domain, we recommend it to the other users in the same community.

To get a subjective evaluation of our recommendation system, we asked for feedback from iLearn users by answering the questions in Fig. 23. We focus here on four evaluation criteria, the recommended resources should be (1) interesting for the user, (2) approachable, (3) adaptable for users if their interests change and (4) divers for users with several interests.

The questionnaire has been answered by 27 of iLearn users. Fig. 24 exhibits the mean of the obtained marks for each criteria. It shows that the users find the recommended resources interesting. Approachability of resources had the best mark from the users, which is explained by the fact that iLearn can detect prerequisites for each learning domain and recommend also resources that build learner's preliminary knowledge before addressing what learner expresses as needs. Fig. 25 shows how our system relies on learning ontology to analyze domains that the user wants to master and detect their prerequisites and recommend the adequate materials. Also Fig. 24 shows that the users are not very satisfied of the recommendation system adaptability to their needs. In fact, most users were neutral about this point (gave mark 3/5), however, we believe that most users did not even have enough time to change their interests and that iLearn will improve its adaptability over time when it gathers more data about users interests. Finally, most users are satisfied about recommendations diversity, which reflects the advantage of relying on a hybrid approach for recommendation (both user's interactions and the interests of members in the same community).

6. Conclusion

This paper develops a fully automatic framework called iLearn that exploits the web semantic in order to improve the learning process, the recommendation quality, and the analysis of social networks. iLearn is dedicated to the learning community that guides tutors to know the students limitations and helps them to understand deeply a specific learning domain by providing the most appropriate orientations, helps, and resources. To accomplish the objectives of iLearn, we developed a social learning platform, learning and emotional ontologies, both of them are scalable and extensible. Further, iLearn models systems resources and inputs as RDF graphs, a standard formalism. The semantic analysis algorithms that detects the different learning communities and produces the appropriate recommendations advance the state-of-the-art crowd intelligence and the anticipatory computing.

Currently and in the near future, we plan to extend this work in different directions especially in crowd intelligence and anticipatory computing oriented fields. First we would like to evaluate how well the provided recommendations fit the learners intentions and the learning objectives. This will help to evaluate the provided recommendations and resources which make iLearn dynamic in terms of reactivity. Further we have to carried out iLearn by studying the social networks relationships that will be presented by weighted arcs according to the nature of their dependences. Furthermore, we target to improve the natural language processing built within iLearn. Another work is in progress which integrates external data sources on the semantic web like DBPedia[3] and using formal methods and theorem provers to prove the soundness and the correctness of iLearn.

References

Aşkar, P., 2011. Social network analysis for e-learning environments. Procedia - Social and Behavioral Sciences 28, 992, world Conference on Educational Technology Researches - 2011.

Cantador, I., Castells, P., 2006. Multilayered semantic social network modeling by ontology-based user profiles clustering: Application to collaborative filtering. In: Managing knowledge in a world of networks. pringer Berlin Heidelberg, pp. 334–349.

Chohra, C., Khaled, A., 2013. Outil d'analyse des réseaux sociaux à base de Web sémantique pour les Environnements d'apprentissage. Master's thesis University of Guelma.

Dráždilová, P., Martinovič, J., Slaninová, K., Snášel, V., 2008. Analysis of relations in eLearning. Web intelligence and intelligent agent technology.

Freeman, L., 1979. Centrality in social networks conceptual clarification. Social Networks 1 (3), 215–239.

Gruber, T., 2008. Collective knowledge systems: Where the social web meets the semantic web. Web Semantics Science Services and Agents on the World Wide Web 6 (1), 4–13.

Halimi, K., Seridi-Bouchelaghem, H., 2015. Semantic web based learning styles identification for social learning environments personalization. Web Intelligence 3, 3–29.

Heflin, J., Hendler, J., 2001. A portrait of the Semantic Web in action. IEEE Intelligent Systems 16, 54–59.

Jetinai, K., 2018. Rule-based reasoning for resource recommendation in personalized e-learning. In: 2018 international conference on information and computer technologies (ICICT). pp. 150–154.

Li, T., Yang, H., He, J., Ai, Y., 2010. A Social Network Analysis methods based on ontology. In: 2010 third international symposium on knowledge acquisition and modeling. pp. 258–261.

Mansur, A.B.F., Yusof, N., 2013. Social learning network analysis model to identify learning patterns using ontology clustering techniques and meaningful learning. Computers & Education 63, 73–86.

Martinez, A., Dimitriadis, Y., Rubia, B., Gómez, E., de la Fuente, P., 2003. Combining qualitative evaluation and social network analysis for the study of classroom social interactions. Computers & Education 41 (4), 353–368.

Nieminen, J., 1974. On the centrality in a graph. Scandinavian Journal of Psychology 15 (4), 332–336.

Obeid, C., Lahoud, I., El Khoury, H., Champin, P.-A., 2018. Ontology-based recommender system in higher education. In: Companion proceedings of the the web conference 2018, WWW '18, international world wide web conferences steering committee. pp. 1031–1034.

Oldakowski, R., Bizer, C., Westphal, D., 2004. RAP: RDF API for PHP. In: In proc. International workshop on interpreted languages. MIT Press.

Y. Ou, C.-q. Zhang, A new multimembership clustering method, Journal of Industrial and Management Optimization 3.

Qi, X., Tang, W., Wu, Y., Guo, G., Fuller, E., Zhang, C.-Q., 2014. Optimal local community detection in social networks based on density drop of subgraphs. Pattern Recognition Letters 36, 46–53.

Scott, J.P., 2000. Social network analysis: A handbook, 2nd ed. Sage Publications Ltd.

Sun, P.G., Sun, X., 2017. Complete graph model for community detection. Physica A: Statistical Mechanics and Its Applications 471, 88–97.

Yusof, N., Mansur, A.B.F., Othman, M.S., 2011. Ontology of Moodle e-Learning System for Social Network Analysis. In: IEEE conference on open systems (ICOS2011). pp. 122–126.

[3] http://wiki.dbpedia.org/.